

Making Repairs in Description Logics More Gentle

Franz Baader
TU Dresden,
Dresden, Germany

Francesco Kriegel
TU Dresden,
Dresden, Germany

Adrian Nuradiansyah
TU Dresden,
Dresden, Germany

Rafael Peñaloza
Free University of Bolzano,
Bolzano, Italy

Abstract

The classical approach for repairing a Description Logic ontology \mathcal{D} in the sense of removing an unwanted consequence α is to delete a minimal number of axioms from \mathcal{D} such that the resulting ontology \mathcal{D}' does not have the consequence α . However, the complete deletion of axioms may be too rough, in the sense that it may also remove consequences that are actually wanted. To alleviate this problem, we propose a more gentle notion of repair in which axioms are not deleted, but only weakened. On the one hand, we investigate general properties of this gentle repair method. On the other hand, we propose and analyze concrete approaches for weakening axioms expressed in the Description Logic \mathcal{EL} .

Introduction

Description logics (DLs) (Baader et al. 2017) are a family of logic-based knowledge representation formalisms, which are employed in various application domains, such as natural language processing, configuration, databases, and biomedical ontologies, but their most notable success so far is the adoption of the DL-based language OWL¹ as standard ontology language for the Semantic Web. As the size of DL-based ontologies grows, tools that support improving the quality of such ontologies become more important. DL reasoners² can be used to detect inconsistencies and to infer other implicit consequences, such as subsumption and instance relationships. However, for the developer of a DL-based ontology, it is often quite hard to understand why a consequence computed by the reasoner actually follows from the knowledge base, and how to repair the ontology in case this consequence is not intended.

Axiom pinpointing (Schlobach and Cornet 2003) was introduced to help developers or users of DL-based ontologies understand the reasons why a certain consequence holds by computing so-called *justifications*, i.e., minimal subsets of the ontology that have the consequence in question. Black-box approaches for computing justifications such as (Schlobach et al. 2007; Kalyanpur et al. 2007;

Baader and Suntisrivaraporn 2008) use repeated calls of existing highly-optimized DL reasoners for this purpose, but it may be necessary to call the reasoner an exponential number of times. In contrast, glass-box approaches such as (Baader and Hollunder 1995; Schlobach and Cornet 2003; Parsia, Sirin, and Kalyanpur 2005; Meyer et al. 2006) compute all justifications by a single run of a modified, but usually less efficient reasoner.

Given all justifications of an unwanted consequence, one can then repair the ontology by removing one axiom from each justification. However, removing complete axioms may also eliminate consequences that are actually wanted. For example, assume that our ontology contains the following terminological axioms:

$$\begin{aligned} Prof &\sqsubseteq \exists employed.Uni \sqcap \exists enrolled.Uni, \\ \exists enrolled.Uni &\sqsubseteq Studi. \end{aligned}$$

These two axioms are a justification for the incorrect consequence that professors are students. While the first axiom is the culprit, removing it completely would also remove the correct consequence that professors are employed by a university. Thus, it would be more appropriate to replace the first axiom by the weaker axiom $Prof \sqsubseteq \exists employed.Uni$. This is the basic idea underlying our *gentle repair approach*. In general, in this approach we weaken one axiom from each justification such that the modified justifications no longer have the consequence.

Approaches for repairing ontologies while keeping more consequences than the classical approach based on completely removing axioms have already been considered in the literature. On the one hand, there are approaches that first modify the given ontology, and then repair this modified ontology using the classical approach. In the work by Horridge, Parsia, and Sattler (2008), a specific syntactic structural transformation is applied to the axioms in an ontology, which replaces them by sets of logically weaker axioms. More recently, Du, Qi, and Fu (2014) have generalized this idea by allowing for different specifications of the structural transformation of axioms. They also introduce a specific structural transformation that is based on specializing left-hand sides and generalizing right-hand sides of axioms in a way that ensures finiteness of the obtained set of axioms. Closer to our gentle repair approach is the one of Lam et al. (2008), which adapts the tracing technique from

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹See <https://www.w3.org/TR/owl2-overview/> for its most recent edition OWL 2.

²See <http://owl.cs.manchester.ac.uk/tools/list-of-reasoners/>.

Baader and Hollunder (1995) to identify not only the axioms that cause a consequence, but also the parts of these axioms that are actively involved in deriving the consequence. This provides them with information for how to weaken these axioms. In (Troquard et al. 2018), repairs are computed by weakening axioms with the help of refinement operators (Lehmann and Hitzler 2010).

In this paper, we introduce a general framework for repairing ontologies based on axiom weakening. This framework is independent of the concrete method employed for weakening axioms and of the concrete ontology language used to write axioms. It only assumes that ontologies are finite sets of axioms, that there is a monotonic consequence operator defining which axioms follow from which, and that weaker axioms have less consequences. However, all our examples will consider ontologies expressed in the light-weight DL \mathcal{EL} . Our first important result is that, in general, the gentle repair approach needs to be iterated, i.e., applying it once does not necessarily remove the consequence. This problem has actually been overlooked in (Lam et al. 2008), which means that their approach does not always yield a repair. Our second result is that at most exponentially many iterations are always sufficient to reach a repair. The authors of (Troquard et al. 2018) had already realized that iteration is needed, but they did not give an example explicitly demonstrating this, and they had no termination proof. Instead of allowing for arbitrary ways of weakening axioms, we then introduce the notion of a *weakening relation*, which restricts the way in which axioms can be weakened. Subsequently, we define conditions on such weakening relations that equip the gentle repair approach with better algorithmic properties if they are satisfied. Finally, we address the task of defining specific weakening relations for the DL \mathcal{EL} . After showing that two quite large such relations do not behave well, we introduce two restricted relations, which are based on generalizing the right-hand sides of axioms semantically or syntactically. Both of them satisfy most of our conditions, but from a complexity point of view the syntactic variant behaves considerably better. Due to space constraints, some of the proofs of our results cannot be given here. They can be found in (Baader et al. 2018).

Basic Definitions

In the first part of this section, we introduce basic notions from DLs to equip us with concrete examples for how ontologies and their axioms may look like. In the second part, we provide basic definitions regarding ontology repair, which are independent of the ontology language these ontologies are written in.

Description Logics

A wide range of DLs of different expressive power have been investigated in the literature. Here, we only introduce the DL \mathcal{EL} , for which reasoning is tractable (Brandt 2004).

Let N_C and N_R be mutually disjoint sets of *concept* and *role names*, respectively. Then \mathcal{EL} *concepts* over these names are constructed from these names using the top concept (\top), conjunction ($C \sqcap D$), and existential restriction ($\exists r.C$).

The *size* of an \mathcal{EL} concept C is the number of occurrences of \top as well as concept and role names in C , and its *role depth* is the maximal nesting of existential restrictions. If S is a finite set of \mathcal{EL} concepts, then we denote the conjunction of these concepts as $\prod S$.

Knowledge is represented using appropriate axioms formulated using concepts, role names and an additional set of individual names N_I . An \mathcal{EL} *axiom* is either a *general concept inclusion* (GCI) of the form $C \sqsubseteq D$ with C, D concepts, or an *assertion* of the form $C(a)$ (*concept assertion*) or $r(a, b)$ (*role assertion*), with $a, b \in N_I, r \in N_R$, and C a concept. A finite set of GCIs is called a *TBox*; a finite set of assertions is an *ABox*. An *ontology* is a finite set of axioms.

The semantics of \mathcal{EL} is defined through *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty set, called the *domain*, and $\cdot^{\mathcal{I}}$ is the *interpretation function*, which maps every $a \in N_I$ to an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, every $A \in N_C$ to a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and every $r \in N_R$ to a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. This function $\cdot^{\mathcal{I}}$ is extended to arbitrary \mathcal{EL} concepts by setting $\top^{\mathcal{I}} := \Delta^{\mathcal{I}}, (C \sqcap D)^{\mathcal{I}} := C^{\mathcal{I}} \cap D^{\mathcal{I}}, (\exists r.C)^{\mathcal{I}} := \{e \in \Delta^{\mathcal{I}} \mid \exists f \in C^{\mathcal{I}}. (e, f) \in r^{\mathcal{I}}\}$.

We say that the interpretation \mathcal{I} *satisfies* the GCI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$; it *satisfies* the assertion $C(a)$ and $r(a, b)$, if $a^{\mathcal{I}} \in C^{\mathcal{I}}$ and $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$, respectively. It is a *model* of the TBox \mathcal{T} , the ABox \mathcal{A} , and the ontology \mathcal{D} , if it satisfies all the axioms in \mathcal{T}, \mathcal{A} , and \mathcal{D} , respectively. Given an ontology \mathcal{D} , and an axiom α , α is a *consequence* of \mathcal{D} (or \mathcal{D} *entails* α) if every model of \mathcal{D} satisfies α . In this case, we write $\mathcal{D} \models \alpha$. The *set of all consequences* of \mathcal{D} is denoted by $Con(\mathcal{D})$. As shown in (Brandt 2004), consequences in \mathcal{EL} can be decided in polynomial time. The two axioms α, β are *equivalent* if $Con(\{\alpha\}) = Con(\{\beta\})$.

A *tautology* is an axiom α such that $\emptyset \models \alpha$, where \emptyset is the ontology that contains no axioms. For example, GCIs of the form $C \sqsubseteq \top$ and $C \sqsubseteq C$, and assertions of the form $\top(a)$ are tautologies. We write $C \sqsubseteq^{\emptyset} D$ to indicate that the GCI $C \sqsubseteq D$ is a tautology. In this case we say that C is *subsumed* by D . The concepts C, D are *equivalent* (written $C \equiv^{\emptyset} D$) if $C \sqsubseteq^{\emptyset} D$ and $D \sqsubseteq^{\emptyset} C$. The concept C is *strictly subsumed* by D (written $C \sqsubset^{\emptyset} D$) if $C \sqsubseteq^{\emptyset} D$ and $C \not\equiv^{\emptyset} D$.

Repairing Ontologies

For the purpose of this subsection and also large parts of the rest of this paper, we leave it open what sort of axioms and ontologies are allowed in general, but we draw our examples from \mathcal{EL} ontologies. We only assume that there is a monotone consequence relation $\mathcal{D} \models \alpha$ between ontologies (i.e., finite sets of axioms) and axioms, and that $Con(\mathcal{D})$ consists of all consequences of \mathcal{D} .

Assume in the following that the ontology $\mathcal{D} = \mathcal{D}_s \cup \mathcal{D}_r$ is the disjoint union of a *static* ontology \mathcal{D}_s and a *refutable* ontology \mathcal{D}_r . When repairing the ontology, only the refutable part may be changed. For example, the static part of the ontology could be a carefully hand-crafted TBox whereas the refutable part is an ABox that is automatically generated from (possibly erroneous) data. It may also make sense to classify parts of a TBox as refutable, for example if the TBox is obtained as a combination of ontologies from

different sources, some of which may be less trustworthy than others. In a privacy application (Cuenca Grau 2010; Baader, Borchmann, and Nuradiansyah 2017), it may be the case that parts of the ontology are publicly known whereas other parts are hidden. In this setting, in order to hide critical information, it only makes sense to change the hidden part of the ontology.

Definition 1. Let $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$ be an ontology consisting of a static and a refutable part, and α an axiom such that $\mathfrak{D} \models \alpha$ and $\mathfrak{D}_s \not\models \alpha$. The ontology \mathfrak{D}' is a repair of \mathfrak{D} w.r.t. α if $\text{Con}(\mathfrak{D}_s \cup \mathfrak{D}') \subseteq \text{Con}(\mathfrak{D}) \setminus \{\alpha\}$.

The repair \mathfrak{D}' is an optimal repair of \mathfrak{D} w.r.t. α if there is no repair \mathfrak{D}'' of \mathfrak{D} w.r.t. α with $\text{Con}(\mathfrak{D}_s \cup \mathfrak{D}') \subset \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}'')$. The repair \mathfrak{D}' is a classical repair of \mathfrak{D} w.r.t. α if $\mathfrak{D}' \subset \mathfrak{D}_r$, and it is an optimal classical repair of \mathfrak{D} w.r.t. α if there is no classical repair \mathfrak{D}'' of \mathfrak{D} w.r.t. α such that $\mathfrak{D}' \subset \mathfrak{D}''$.

The condition $\mathfrak{D}_s \not\models \alpha$ ensures that \mathfrak{D} does have a repair w.r.t. α since obviously the empty ontology \emptyset is such a repair. In general, optimal repairs need not exist.

Proposition 2. There is an \mathcal{EL} ontology $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$ and an \mathcal{EL} axiom α such that \mathfrak{D} does not have an optimal repair w.r.t. α .

Proof. We set $\alpha := A(a)$, $\mathfrak{D}_s := \mathcal{T}$, and $\mathfrak{D}_r := \mathcal{A}$ where

$$\mathcal{T} := \{A \sqsubseteq \exists r.A, \exists r.A \sqsubseteq A\} \text{ and } \mathcal{A} := \{A(a)\}.$$

To show that there is no optimal repair of \mathfrak{D} w.r.t. α , we consider an arbitrary repair \mathfrak{D}' and show that it cannot be optimal. Thus, let \mathfrak{D}' be such that

$$\text{Con}(\mathcal{T} \cup \mathfrak{D}') \subseteq \text{Con}(\mathfrak{D}) \setminus \{A(a)\}.$$

Without loss of generality we assume that \mathfrak{D}' contains assertions only: if \mathfrak{D}' contains a GCI that does not follow from \mathcal{T} , then $\text{Con}(\mathcal{T} \cup \mathfrak{D}') \not\subseteq \text{Con}(\mathfrak{D})$. This is an easy consequence of the fact that, in \mathcal{EL} , a GCI follows from a TBox together with an ABox iff it follows from the TBox alone. It is also easy to see that \mathfrak{D}' cannot contain role assertions since no such assertions are entailed by \mathfrak{D} . In addition, concept assertions following from $\mathcal{T} \cup \mathfrak{D}'$ must have a specific form.

Claim: If the assertion $C(a)$ is in $\text{Con}(\mathcal{T} \cup \mathfrak{D}')$, then C does not contain A .

Proof of claim. By induction on the role depth n of C .

Base case: If $n = 0$ and A is contained in C , then A is a conjunct of C and thus $C(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$ implies $A(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$, which is a contradiction.

Step case: If $n > 0$ and A occurs at role depth n in C , then $C(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$ implies that there are roles r_1, \dots, r_n such that $(\exists r_1 \dots \exists r_n.A)(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$. Since $\text{Con}(\mathcal{T} \cup \mathfrak{D}') \subseteq \text{Con}(\mathfrak{D})$, this can only be the case if $r_1 = \dots = r_n = r$ since \mathfrak{D} clearly has models in which all roles different from r are empty. Since \mathcal{T} contains the GCI $\exists r.A \sqsubseteq A$ and $r_n = r$, $(\exists r_1 \dots \exists r_n.A)(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$ implies $(\exists r_1 \dots \exists r_{n-1}.A)(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$. Induction now yields that this is not possible, which completes the proof of the claim.

Furthermore, as argued in the proof of the claim, any assertion belonging to $\text{Con}(\mathfrak{D})$ cannot contain roles other than

r . The same is true for concept names different from A . Consequently, all assertions $C(a) \in \text{Con}(\mathcal{T} \cup \mathfrak{D}')$ are such that C is built using r and \top only. Any such concept C is equivalent to a concept of the form $(\exists r.)^n \top$.

Since \mathfrak{D}' is finite, there is a maximal n_0 such that $((\exists r.)^{n_0} \top)(a) \in \mathfrak{D}'$, but $((\exists r.)^n \top)(a) \notin \mathfrak{D}'$ for all $n > n_0$. Since $(\exists r.)^n \top \sqsubseteq (\exists r.)^m \top$ if $m \leq n$, we can assume without loss of generality that $\mathfrak{D}' = \{((\exists r.)^{n_0} \top)(a)\}$. It is now easy to show that $((\exists r.)^n \top)(a) \notin \text{Con}(\mathcal{T} \cup \mathfrak{D}')$ if $n > n_0$. Consequently, if we choose n such that $n > n_0$ and define $\mathfrak{D}'' := \{((\exists r.)^n \top)(a)\}$, then $\text{Con}(\mathcal{T} \cup \mathfrak{D}') \subset \text{Con}(\mathcal{T} \cup \mathfrak{D}'')$. In addition, $\text{Con}(\mathcal{T} \cup \mathfrak{D}'') \subseteq \text{Con}(\mathfrak{D}) \setminus \{A(a)\}$, i.e., \mathfrak{D}'' is a repair. This shows that \mathfrak{D}' is not optimal. Since we have chosen \mathfrak{D}' to be an arbitrary repair, this shows that there cannot be an optimal repair. \square

In contrast, optimal classical repairs always exist. One approach for computing such a repair uses justifications and hitting sets (Reiter 1987).

Definition 3. Let $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$ be an ontology and α an axiom such that $\mathfrak{D} \models \alpha$ and $\mathfrak{D}_s \not\models \alpha$. A justification for α in \mathfrak{D} is a minimal subset J of \mathfrak{D}_r such that $\mathfrak{D}_s \cup J \models \alpha$. Given justifications J_1, \dots, J_k for α in \mathfrak{D} , a hitting set of these justifications is a set H of axioms such that $H \cap J_i \neq \emptyset$ for $i = 1, \dots, k$. This hitting set is minimal if there is no other hitting set strictly contained in it.

Note that the condition $\mathfrak{D}_s \not\models \alpha$ implies that justifications are non-empty. Consequently, hitting sets and thus minimal hitting sets always exist.

The algorithm for computing an optimal classical repair of \mathfrak{D} w.r.t. α proceeds in two steps: (i) compute all justifications J_1, \dots, J_k for α in \mathfrak{D} ; and then (ii) compute a minimal hitting set H of J_1, \dots, J_k and remove the elements of H from \mathfrak{D}_r , i.e., output $\mathfrak{D}' = \mathfrak{D}_r \setminus H$.

It is not hard to see that, independently of the choice of the hitting set, this algorithm produces an optimal classical repair. Conversely, all optimal classical repairs can be generated this way by going through all hitting sets.

Gentle Repairs

Instead of removing axioms completely, as in the case of classical repairs, gentle repairs replace them by weaker axioms.

Definition 4. Let β, γ be axioms. Then γ is weaker than β if $\text{Con}(\{\gamma\}) \subset \text{Con}(\{\beta\})$.

Alternatively, we could have introduced the notion of weaker w.r.t the strict part of the ontology, by requiring that $\text{Con}(\mathfrak{D}_s \cup \{\gamma\}) \subset \text{Con}(\mathfrak{D}_s \cup \{\beta\})$.³ In this paper, we do not consider this alternative definition, although most of the results in this section would also hold w.r.t. it (e.g., Theorem 6). The difference between the two definitions is, however, relevant in the next section, where we consider concrete approaches for how to weaken axioms. If the whole ontology is refutable, there is of course no difference between the two definitions.

³Defining weaker w.r.t the whole ontology \mathfrak{D} does not make sense since this ontology is possibly erroneous.

Obviously, the *weaker-than* relation from Definition 4 is transitive, i.e., if α is weaker than β and β is weaker than γ , then α is also weaker than γ . In addition, a tautology is always weaker than a non-tautology. Replacing an axiom by a tautology is obviously the same as removing this axiom. We assume in the following that there exist tautological axioms, which is obviously true for \mathcal{EL} .

Gentle repair algorithm: we still compute all justifications J_1, \dots, J_k for α in \mathfrak{D} and a minimal hitting set H of J_1, \dots, J_k , but instead of removing the elements of H from \mathfrak{D}_r , we replace them by weaker axioms. More precisely, if $\beta \in H$ and $J_{i_1}, \dots, J_{i_\ell}$ are all the justifications containing β , then replace β by a weaker axiom γ such that

$$\mathfrak{D}_s \cup (J_{i_j} \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha \text{ for } j = 1, \dots, \ell. \quad (1)$$

Note that such a weaker axiom γ always exists. In fact, we can choose a tautology as axiom γ . If γ is a tautology, then replacing β by γ is the same as removing β . Thus, we have $\mathfrak{D}_s \cup (J_{i_j} \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$ due to the minimality of J_{i_j} . In addition, minimality of J_{i_j} also implies that β is not a tautology since otherwise $\mathfrak{D}_s \cup (J_{i_j} \setminus \{\beta\})$ would also have the consequence α . In general, different choices of γ yield different runs of the algorithm.

In principle, the algorithm could always use a tautology γ , but then this run would produce a classical repair. To obtain more gentle repairs, the algorithm needs to use a strategy that chooses stronger axioms (i.e., axioms γ that are less weak than tautologies) if possible. In contrast to what is claimed in the literature (Lam et al. 2008), this approach does not necessarily yield a repair.

Lemma 5. *Let \mathfrak{D}' be the ontology obtained from \mathfrak{D}_r by replacing all elements of the hitting set by weaker ones such that (1) is satisfied. Then $\text{Con}(\mathfrak{D}_s \cup \mathfrak{D}') \subseteq \text{Con}(\mathfrak{D})$, but in general we may still have $\alpha \in \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}')$.*

Proof. The definition of “weaker” (Definition 4) obviously implies that $\text{Con}(\mathfrak{D}_s \cup \mathfrak{D}') \subseteq \text{Con}(\mathfrak{D})$.

We give an example where this approach does not produce a repair. Let $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$ for $\mathfrak{D}_s = \mathcal{A}$ and $\mathfrak{D}_r = \mathcal{T}$, where $\mathcal{A} = \{A(a), r(a, a)\}$, and $\mathcal{T} = \{A \sqsubseteq B, \exists r.B \sqsubseteq B\}$, and let α be the consequence $B(a)$. The only justification for α is $\{\beta = A \sqsubseteq B\}$. The weakening $\gamma = \exists r.A \sqsubseteq \exists r.B$ of β satisfies (1), but the resulting ontology $\mathfrak{D}_s \cup (\mathfrak{D}_r \setminus \{\beta\}) \cup \{\gamma\}$ still implies $B(a)$. \square

The example from the previous proof shows that applying the gentle repair approach only once may not lead to a repair. For this reason, we need to *iterate this approach*, i.e., if the resulting ontology $\mathfrak{D}_s \cup \mathfrak{D}'$ still has α as a consequence, we again compute all justifications and a hitting set for them, and then replace the elements of the hitting set with weaker axioms as described above. This is iterated until a repair is reached. We can show that this iteration indeed always terminates after finitely many steps with a repair.

Theorem 6. *Let $\mathfrak{D}^{(0)} = \mathfrak{D}_s^{(0)} \cup \mathfrak{D}_r^{(0)}$ be a finite ontology and α an axiom such that $\mathfrak{D}^{(0)} \models \alpha$ and $\mathfrak{D}_s^{(0)} \not\models \alpha$. Applied to $\mathfrak{D}^{(0)}$ and α , the iterative algorithm described above stops*

after a finite number of iterations that is at most exponential in the cardinality of $\mathfrak{D}_r^{(0)}$, and yields as output an ontology that is a repair of $\mathfrak{D}_s^{(0)}$ w.r.t. α .

Proof. Assume that $\mathfrak{D}_r^{(0)}$ contains n axioms, and that there is an infinite run R of the algorithm on input $\mathfrak{D}^{(0)}$ and α . Take a bijection ℓ_0 between $\mathfrak{D}_r^{(0)}$ and $\{1, \dots, n\}$ that assigns unique labels to axioms. Whenever we weaken an axiom during a step of the run, the new weaker axiom inherits the label of the original axiom. Thus, we have bijections $\ell_i : \mathfrak{D}_r^{(i)} \rightarrow \{1, \dots, n\}$ for all ontologies $\mathfrak{D}_r^{(i)}$ considered during the run R of the algorithm. For $i \geq 0$ we define

$$S_i := \{K \subseteq \{1, \dots, n\} \mid \mathfrak{D}_s \cup \{\beta \in \mathfrak{D}_r^{(i)} \mid \ell_i(\beta) \in K\} \models \alpha\},$$

i.e., S_i contains all sets of indices such that the corresponding subset of $\mathfrak{D}_r^{(i)}$ together with \mathfrak{D}_s has the consequence α .

We claim that $S_{i+1} \subset S_i$. Note that $S_{i+1} \subseteq S_i$ is an immediate consequence of the fact that $\ell_i(\gamma) = j = \ell_{i+1}(\gamma')$ implies that $\gamma = \gamma'$ or γ' is weaker than γ . Thus, it remains to show that the inclusion is strict. This follows from the following observations. Since the algorithm does not terminate with the ontology $\mathfrak{D}_r^{(i)}$, we still have $\mathfrak{D}_s \cup \mathfrak{D}_r^{(i)} \models \alpha$, and thus there is at least one justification $\emptyset \subset J \subseteq \mathfrak{D}_r^{(i)}$. Consequently, the hitting set H used in this step of the algorithm contains an element β of $\mathfrak{D}_r^{(i)}$. When going from $\mathfrak{D}_r^{(i)}$ to $\mathfrak{D}_r^{(i+1)}$, β is replaced by a weaker axiom β' such that $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\beta'\} \not\models \alpha$. But then the set $\{\ell_i(\gamma) \mid \gamma \in J\}$ belongs to S_i , but not to S_{i+1} .

Since S_0 contains at most exponentially many sets, the strict inclusion $S_{i+1} \subset S_i$ can happen at most exponentially often, which contradicts our assumption that there is an infinite run R of the algorithm on input $\mathfrak{D}^{(0)}$ and α . This shows termination after exponentially many steps. However, if the algorithm terminates with output $\mathfrak{D}_r^{(i)}$, then $\mathfrak{D}_s \cup \mathfrak{D}_r^{(i)} \not\models \alpha$; otherwise, it would be possible to weaken $\mathfrak{D}_r^{(i)}$ into $\mathfrak{D}_r^{(i+1)}$ since it would always be possible to replace the elements of a hitting set by tautologies, i.e., perform a classical repair. \square

When computing a classical repair, considering all justifications and then removing a minimal hitting set of these justifications guarantees that one immediately obtains a repair. We have seen in the proof of Lemma 5 that with our gentle repair approach this need not be the case. Nevertheless, we were able to show that, after a finite number of iterations of the approach, we obtain a repair. The proof of termination actually shows that it is sufficient to weaken only one axiom of one justification such that the resulting set is no longer a justification. This motivates the following modification of our approach:

Modified gentle repair algorithm: compute one justifications J for α in \mathfrak{D} and choose an axiom $\beta \in J$. Replace β by a weaker axiom γ such that

$$\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha. \quad (2)$$

Clearly, one needs to iterate this approach, but it is easy to see that the termination argument used in the proof of Proposition 6 also applies here.

Corollary 7. *Let $\mathfrak{D}^{(0)} = \mathfrak{D}_s^{(0)} \cup \mathfrak{D}_r^{(0)}$ be a finite ontology and α an axiom such that $\mathfrak{D}^{(0)} \models \alpha$ and $\mathfrak{D}_s^{(0)} \not\models \alpha$. Applied to $\mathfrak{D}^{(0)}$ and α , the modified iterative algorithm stops after a finite number of iterations that is at most exponential in the cardinality of $\mathfrak{D}_r^{(0)}$, and yields as output an ontology that is a repair of $\mathfrak{D}_s^{(0)}$ w.r.t. α .*

An important advantage of this modified approach is that the complexity of a single iteration step may decrease considerably. For example, for the DL \mathcal{EL} , a single justification can be computed in polynomial time, while computing all justifications may take exponential time (Baader, Peñaloza, and Suntisrivaraporn 2007). In addition, to compute a minimal hitting set one needs to solve an NP-complete problem (Garey and Johnson 1979) whereas choosing one axiom from a single justification is easy. However, as usual, there is no free lunch: we can show that the modified gentle repair algorithm may indeed need exponentially many iteration steps.⁴

Proposition 8. *There exists a sequence of \mathcal{EL} ontologies $\mathfrak{D}^{(n)} = \mathfrak{D}_s^{(n)} \cup \mathfrak{D}_r^{(n)}$ with $\mathfrak{D}_s^{(n)} = \emptyset$ and an \mathcal{EL} axiom α such that the modified gentle repair algorithm applied to $\mathfrak{D}^{(n)}$ and α has a run with exponentially many iterations in the size of $\mathfrak{D}^{(n)}$.*

Proof. Let $I^{(n)} = \{P_i, Q_i \mid 1 \leq i \leq n\}$, $n \geq 1$ be a set of concept names, and define $\mathfrak{D}^{(n)} := \mathfrak{D}_r^{(n)} := \mathcal{T}_1^{(n)} \cup \mathcal{T}_2^{(n)}$, where

$$\begin{aligned} \mathcal{T}_1^{(n)} &:= \{A \sqsubseteq \exists r. \prod I^{(n)}, \exists r. (P_n \sqcap Q_n) \sqsubseteq B\} \cup \\ &\quad \{P_i \sqcap Q_i \sqsubseteq P_{i+1}, P_i \sqcap Q_i \sqsubseteq Q_{i+1} \mid 1 \leq i < n\}, \\ \mathcal{T}_2^{(n)} &:= \{\exists r. (X \sqcap Y) \sqsubseteq D_{XY}, D_{XY} \sqcap X \sqsubseteq Y \mid \\ &\quad X \in \{P_i, Q_i\}, Y \in \{P_{i+1}, Q_{i+1}\}, 1 \leq i < n\} \cup \\ &\quad \{\exists r. P_1 \sqsubseteq P_1, \exists r. Q_1 \sqsubseteq Q_1, P_n \sqsubseteq B, Q_n \sqsubseteq B\}. \end{aligned}$$

It is easy to see that the size of $\mathfrak{D}^{(n)}$ is polynomial in n and that $\mathfrak{D}^{(n)} \models A \sqsubseteq B$. Suppose that we want to get rid of this consequence using the modified gentle repair approach. First, we can find the justification

$$\{A \sqsubseteq \exists r. \prod I^{(n)}, \exists r. (P_n \sqcap Q_n) \sqsubseteq B\}.$$

We repair it by weakening the first axiom to

$$\gamma := A \sqsubseteq \exists r. \prod (I^{(n)} \setminus \{P_n\}) \sqcap \exists r. \prod (I^{(n)} \setminus \{Q_n\}).$$

At this point, we can find a justification that uses γ and $P_{n-1} \sqcap Q_{n-1} \sqsubseteq P_n$. We further weaken γ to

$$\begin{aligned} A \sqsubseteq \exists r. \prod (I^{(n)} \setminus \{P_n, P_{n-1}\}) \sqcap \\ \exists r. \prod (I^{(n)} \setminus \{P_n, Q_{n-1}\}) \sqcap \exists r. \prod (I^{(n)} \setminus \{Q_n\}). \end{aligned}$$

⁴It is not clear yet whether this is also the case for the unmodified gentle repair algorithm.

Repeating this approach, after $2n$ weakenings we have only changed the first axiom, weakening it to the axiom

$$A \sqsubseteq \prod_{X_i \in \{P_i, Q_i\}, 1 \leq i \leq n} \exists r. (X_1 \sqcap \dots \sqcap X_n), \quad (3)$$

whose right-hand side is a conjunction with 2^n conjuncts, each of them representing a possible choice of P_i or Q_i at every location i , $1 \leq i \leq n$.

So far, we have just considered axioms from $\mathcal{T}_1^{(n)}$. Taking also axioms from $\mathcal{T}_2^{(n)}$ into account, we obtain for every conjunct $\exists r. (X_1 \sqcap \dots \sqcap X_n)$ in axiom (3) a justification for $A \sqsubseteq B$ that consists of (3) and the axioms

$$\begin{aligned} \{\exists r. X_1 \sqsubseteq X_1, X_n \sqsubseteq B\} \cup \\ \{\exists r. (X_i \sqcap X_{i+1}) \sqsubseteq D_{X_i X_{i+1}}, \\ D_{X_i X_{i+1}} \sqcap X_i \sqsubseteq X_{i+1} \mid 1 \leq i < n\}. \end{aligned}$$

This justification can be removed by weakening (3) further by deleting one concept name appearing in the conjunct. The justifications for other conjuncts are not influenced by this modification. Thus, we can repeat this for each of the exponentially many conjuncts, which shows that overall we have exponentially many iterations of the modified gentle repair algorithm in this run. \square

Weakening relations

To obtain better bounds on the number of iterations of our algorithms, we restrict the way in which axioms can be weakened. Before introducing concrete approaches for how to do this for \mathcal{EL} axioms in the next section, we investigate such restricted weakening relations in a more abstract setting.

Definition 9. *Given a pre-order \succ (i.e., an irreflexive and transitive binary relation) on axioms, we say that it is*

- a weakening relation if $\beta \succ \gamma$ implies $\text{Con}(\{\gamma\}) \subseteq \text{Con}(\{\beta\})$;
- bounded (linear, polynomial) if, for every axiom α , there is a (linear, polynomial) bound $b(\alpha)$ on the length of all \succ -chains issuing from α ;
- complete if, for any axiom β that is not a tautology, there is a tautology γ such that $\beta \succ \gamma$.

If we use a linear (polynomial) and complete weakening relation, then termination with a repair is guaranteed after a linear (polynomial) number of iterations.

Proposition 10. *Let \succ be a linear (polynomial) and complete weakening relation. If in the (modified) gentle repair algorithm we have $\beta \succ \gamma$ whenever β is replaced by γ , then the algorithm stops after a linear (polynomial) number of iterations and yields as output an ontology that is a repair of $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$ w.r.t. α .*

Proof. For every axiom β in \mathfrak{D}_r we consider the length of the longest \succ -chain issuing from it, and then sum up these numbers over all axioms in \mathfrak{D}_r . The resulting number is linearly (polynomially) bounded by the size of the ontology (assuming that this size is given as sum of the sizes of all its axioms). Call this number the *chain-size* of the ontology. Obviously, if β is replaced by β' with $\beta \succ \beta'$, then the length

of the longest \succ -chain issuing from β' is smaller than the length of the longest \succ -chain issuing from β . Consequently, if $\mathfrak{D}_r^{(i+1)}$ is obtained from $\mathfrak{D}_r^{(i)}$ in the i -th iteration of the algorithm, then the chain-size of $\mathfrak{D}_r^{(i)}$ is strictly larger than the chain-size of $\mathfrak{D}_r^{(i+1)}$. This implies that there can be only linearly (polynomially) many iterations.

Consider a terminating run of the algorithm that produces the sequence of ontologies $\mathfrak{D}_r = \mathfrak{D}_r^{(0)}, \mathfrak{D}_r^{(1)}, \dots, \mathfrak{D}_r^{(n)}$. Then, since \succ is a weakening relation, we have that $\text{Con}(\mathfrak{D}_s \cup \mathfrak{D}_r) \supseteq \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}_r^{(1)}) \supseteq \dots \supseteq \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}_r^{(n)})$. If the algorithm terminated because $\alpha \notin \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}_r^{(n)})$, then $\mathfrak{D}_r^{(n)}$ is a repair of \mathfrak{D} w.r.t. α . Otherwise, the only reason for termination is that, although $\alpha \in \text{Con}(\mathfrak{D}_s \cup \mathfrak{D}_r^{(n)})$, the algorithm cannot generate a new ontology $\mathfrak{D}_r^{(n+1)}$. In the unmodified gentle repair approach this means that there is an axiom β in the hitting set H such that there is no axiom γ with $\beta \succ \gamma$ satisfying (1). However, using a tautology as the axiom γ actually allows us to satisfy the condition (1). Thus, completeness of \succ implies that this reason for termination without success cannot occur. An analogous argument can be used for the modified gentle repair approach. \square

When describing our (modified) gentle repair algorithm, we have said that the chosen axiom β needs to be replaced by a weaker axiom γ such that (1) or (2) holds. But we have not said how such an axiom γ can be found. This of course depends on which ontology language and which weakening relation is used. In the abstract setting of this section, we assume that an ‘‘oracle’’ provides us with a weaker axiom.

Definition 11. *Let \succ be a weakening relation. An oracle for \succ is a computable function W that, given an axiom β that is not \succ -minimal, provides us with an axiom $W(\beta)$ such that $\beta \succ W(\beta)$. For \succ -minimal axioms β we assume that $W(\beta) = \beta$.*

If the weakening relation is complete and well-founded (i.e., there is no infinite descending \succ -chain $\beta_1 \succ \beta_2 \succ \beta_3 \succ \dots$), we can effectively find an axiom γ such that (1) or (2) holds. We show this formally only for (2), but condition (1) can be treated similarly.

Lemma 12. *Assume that J is a justification for the consequence α , and $\beta \in J$. If \succ is a well-founded and complete weakening relation and W is an oracle for \succ , then there is an $n \geq 1$ such that (2) holds for $\gamma = W^n(\beta)$. If \succ is additionally linear (polynomial), then n is linear (polynomial) in the size of β .*

Proof. Well-foundedness implies that the \succ -chain $\beta \succ W(\beta) \succ W(W(\beta)) \succ \dots$ is finite; thus there is an n such that $W^{n+1}(\beta) = W^n(\beta)$, i.e., $W^n(\beta)$ is \succ -minimal. Since \succ is complete, this implies that $W^n(\beta)$ is a tautology. Minimality of the justification J yields $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{W^n(\beta)\} \not\models \alpha$. Linearity (polynomiality) of \succ ensures that the length of the \succ -chain $\beta \succ W(\beta) \succ W(W(\beta)) \succ \dots$ is linearly (polynomially) bounded by the size of β . \square

Thus, to find an axiom γ satisfying (1) or (2), we iteratively apply W to β until an axiom satisfying the required

property is found. The proof of Lemma 12 shows that at the latest this is the case when a tautology is reached, but of course the property may already be satisfied before that by a non-tautological axiom $W^i(\beta)$.

In order to weaken axioms as gently as possible, W should realize small weakening steps. The smallest such step is one where there is no step in between.

Definition 13. *Let \succ be a pre-order. The one-step relation⁵ induced by \succ is defined as*

$$\succ_1 := \{(\beta, \gamma) \in \succ \mid \text{there is no } \delta \text{ such that } \beta \succ \delta \succ \gamma\}.$$

We say that \succ_1 covers \succ if its transitive closure is again \succ , i.e., $\succ_1^+ = \succ$. In this case we also say that \succ is one-step generated.

If \succ is one-step generated, then every weaker element can be reached by a finite sequence of one-step weakenings, i.e., if $\beta \succ \gamma$, then there are finitely many elements $\delta_0, \dots, \delta_n$ ($n \geq 1$) such that $\beta = \delta_0 \succ_1 \delta_1 \succ_1 \dots \succ_1 \delta_n = \gamma$. This leads us to the following characterization of pre-orders that are not one-step generated.

Lemma 14. *The pre-order \succ is not one-step generated iff there is a pair of comparable elements $\beta \succ \gamma$ such that every finite chain $\beta = \delta_0 \succ \delta_1 \succ \dots \succ \delta_n = \gamma$ can be refined in the sense that there is an i , $0 \leq i < n$, and an element δ such that $\delta_i \succ \delta \succ \delta_{i+1}$.*

If $\beta \succ \gamma$ are such that any finite chain between them can be refined, then obviously there cannot be an upper bound on the length of the chains issuing from β . Thus, Lemma 14 implies the following result.

Proposition 15. *If \succ is bounded, then it is one-step generated.*

The following example shows that well-founded pre-orders need not be one-step generated.

Example 16. *Consider the pre-order \succ on the set*

$$P := \{\beta\} \cup \{\delta_i \mid i \geq 0\},$$

where $\beta \succ \delta_i$ for all $i \geq 0$, and $\delta_i \succ \delta_j$ iff $i > j$. It is easy to see that \succ is well-founded and $\succ_1 = \{(\delta_{i+1}, \delta_i) \mid i \geq 0\}$. Consequently, \succ_1^+ contains none of the tuples (β, δ_i) for $i \geq 0$, which shows that \succ_1 does not cover \succ . In particular, any finite chain between β and δ_i can be refined.

If we add elements γ_i ($i \geq 0$) with $\beta \succ \gamma_i \succ \delta_i$ to this pre-order, then it becomes one-step generated.

One-step generated weakening relations allow us to find maximally strong weakenings satisfying (1) or (2). Again, we consider only condition (2), but all definitions and results can be adapted to deal with (1) as well.

Definition 17. *Let J be a justification for the consequence α , and $\beta \in J$. We say that γ is a maximally strong weakening of β in J if $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$, but $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\delta\} \models \alpha$ for all δ with $\beta \succ \delta \succ \gamma$.*

In general, maximally strong weakenings need not exist. As an example, assume that the pre-order introduced in Example 16 (without the added axioms γ_i) is a weakening relation on axioms, and assume that $J = \{\beta\}$ and that none

⁵This is sometimes also called the transitive reduction of \succ .

of the axioms δ_i have the consequence. Obviously, in this situation there is no maximally strong weakening of β in J .

Next, we introduce conditions under which maximally strong weakenings always exist, and can also be computed. We say that the one-step generated weakening relation \succ is *effectively finitely (linearly) branching* if for every axiom β the cardinality of the set $\{\gamma \mid \beta \succ_1 \gamma\}$ is finite (linear in the size of β) and this set can effectively be computed.

Proposition 18. *Let \succ be a well-founded, one-step generated, and effectively finitely branching weakening relation and assume that the consequence relation \models is decidable. Then all maximally strong weakenings of an axiom in a justification can effectively be computed.*

Proof. Let J be a justification for the consequence α , and $\beta \in J$. Since \succ is well-founded, one-step generated, and finitely branching, König's Lemma implies that there are only finitely many γ such that $\beta \succ \gamma$, and all these γ can be reached by following \succ_1 . Thus, by a breadth-first search, we can compute the set of all γ such that there is a path $\beta \succ_1 \delta_1 \succ_1 \dots \succ_1 \delta_n \succ_1 \gamma$ with $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$, but $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\delta_i\} \models \alpha$ for all $i, 1 \leq i \leq n$. If this set still contains elements that are comparable w.r.t. \succ (i.e., there is a \succ_1 -path between them), then we remove the weaker elements. It is easy to see that the remaining set consists of all maximally strong weakenings of β in J . \square

Note that the additional removal of weaker elements in the last proof is really necessary. Assume that $\beta \succ_1 \delta_1 \succ_1 \gamma$ and $\beta \succ_1 \delta_2 \succ_1 \gamma$, and $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$, $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\delta_1\} \models \alpha$, but $\mathfrak{D}_s \cup (J \setminus \{\beta\}) \cup \{\delta_2\} \not\models \alpha$. Then both δ_2 and γ belong to the set computed in the breadth-first search, but only δ_2 is a maximally strong weakening (see Example 24, where it is shown that this situation can really occur when repairing \mathcal{EL} ontologies).

In particular, this also means that iterated application of a one-step oracle, i.e., an oracle W satisfying $\beta \succ_1 W(\beta)$, does not necessarily yield a maximally strong weakening.

Weakening Relations for \mathcal{EL} Axioms

In this section, we restrict the attention to \mathcal{EL} ontologies, but some of our approaches and results could also be transferred to other DLs. We start with observing that weakening relations for \mathcal{EL} axioms need not be one-step generated.

Proposition 19. *If we define $\beta \succ^g \gamma$ if $\text{Con}(\gamma) \subset \text{Con}(\beta)$, then \succ^g is a weakening relation on \mathcal{EL} axioms that is not one-step generated.*

Proof. It is obvious that \succ^g is a weakening relation.⁶ To see that it is not one-step generated, consider a GCI β that is not a tautology and an arbitrary tautology γ . Then we have $\beta \succ \gamma$. Let $\beta = \delta_0 \succ^g \delta_1 \succ^g \dots \succ^g \delta_n = \gamma$ be a finite chain leading from β to γ . Then δ_{n-1} must be a GCI that is not a tautology. Assume that $\delta_{n-1} = C \sqsubseteq D$. Then $\delta := \exists r.C \sqsubseteq \exists r.D$ satisfies $\delta_{n-1} \succ^g \delta \succ^g \gamma$. By Lemma 14, this shows that \succ is not one-step generated. \square

⁶In fact, it is the greatest one w.r.t. set inclusion.

Our main idea for obtaining more well-behaved weakening relations is to weaken a GCI $C \sqsubseteq D$ by generalizing the right-hand side D and/or by specializing the left-hand side C . Similarly, a concept assertion $D(a)$ can be weakened by generalizing D . For role assertions we can use as weakening an arbitrary tautological axiom, but will no longer consider them explicitly in the following.

Proposition 20. *If we define*

$$\begin{aligned} C \sqsubseteq D \succ^s C' \sqsubseteq D' & \text{ if } C' \sqsubseteq^\emptyset C, D \sqsubseteq^\emptyset D' \text{ and} \\ & \{C' \sqsubseteq D'\} \not\models C \sqsubseteq D, \\ D(a) \succ^s D'(a) & \text{ if } D \sqsubseteq^\emptyset D', \end{aligned}$$

then \succ^s is a complete weakening relation.

Proof. To prove that \succ^s is a weakening relation we must show that $\beta \succ^s \gamma$ implies $\text{Con}(\{\gamma\}) \subset \text{Con}(\{\beta\})$. If $C' \sqsubseteq^\emptyset C$ and $D \sqsubseteq^\emptyset D'$, then $\text{Con}(\{C' \sqsubseteq D'\}) \subseteq \text{Con}(\{C \sqsubseteq D\})$ and $\text{Con}(\{a : D'\}) \subseteq \text{Con}(\{a : D\})$. The second inclusion is strict iff $D \sqsubseteq^\emptyset D'$. However, for the first inclusion to be strict, $C' \sqsubseteq^\emptyset C$ or $D \sqsubseteq^\emptyset D'$ is a necessary condition, but it is not sufficient. This is why we explicitly require $\{C' \sqsubseteq D'\} \not\models C \sqsubseteq D$, which yields strictness of the inclusion. Completeness is trivial due to the availability of all tautologies of the form $C \sqsubseteq \top$ and $\top(a)$. \square

To see why, for example, $D \sqsubseteq^\emptyset D'$ does not imply $\text{Con}(\{C \sqsubseteq D'\}) \subset \text{Con}(\{C \sqsubseteq D\})$, consider the following example: we have $A \sqcap \exists r.A \sqsubseteq^\emptyset \exists r.A$, but $\text{Con}(\{A \sqsubseteq \exists r.A\}) = \text{Con}(\{A \sqsubseteq A \sqcap \exists r.A\})$.

Unfortunately, the weakening relation \succ^s introduced in Proposition 20 is *not well-founded* since left-hand sides can be specialized indefinitely. For example, we have $\top \sqsubseteq A \succ^s \exists r.\top \sqsubseteq A \succ^s \exists r.\exists r.\top \sqsubseteq A \succ^s \dots$. To avoid this problem, we now restrict the attention to sub-relations of \succ^s that only generalize the right-hand sides of GCIs. We will not consider concept assertions, but they can be treated similarly.

Generalizing the Right-Hand Sides of GCIs

We define

$$C \sqsubseteq D \succ^{sub} C' \sqsubseteq D' \text{ if } C' = C \text{ and } C \sqsubseteq D \succ^s C' \sqsubseteq D'.$$

Theorem 21. *The relation \succ^{sub} on \mathcal{EL} axioms is a well-founded, complete, and one-step generated weakening relation, but it is not polynomial.*

Proof. Proposition 20 implies that \succ^{sub} is a weakening relation and completeness follows from the fact that $C \sqsubseteq D \succ^{sub} C \sqsubseteq \top$ whenever $C \sqsubseteq D$ is not a tautology. In \mathcal{EL} , the inverse subsumption relation is well-founded, i.e., there cannot be an infinite sequence $C_0 \sqsubseteq^\emptyset C_1 \sqsubseteq^\emptyset C_2 \sqsubseteq^\emptyset \dots$ of \mathcal{EL} concepts. Looking at the proof of this result given in (Baader and Morawska 2010), one sees that it actually shows that \sqsubseteq^\emptyset is bounded. This implies that \succ^{sub} is bounded as well, and thus one-step generated by Proposition 15.

It remains to show that \succ^{sub} is not polynomial. Let $n \geq 1$ and $N_n := \{A_1, \dots, A_{2n}\}$ be a set of $2n$ distinct concept names. Then we have

$$\exists r. \prod N_n \sqsubseteq^\emptyset \prod_{X \subseteq N_n, |X|=n} \exists r. \prod X.$$

Note that the size of $\exists r. \sqcap N_n$ is linear in n , but that the conjunction on the right-hand side of this strict subsumption consists of exponentially many concepts $\exists r. \sqcap X$ that are incomparable w.r.t. subsumption. Consequently, by removing one conjunct at a time, we can generate an ascending chain w.r.t. \sqsubseteq^\emptyset of \mathcal{EL} concepts whose length is exponential in n . Using these concepts as right-hand sides of GCIs with left-hand side B for a concept name $B \notin N_n$, we obtain an exponentially long descending chain w.r.t. \succ^{sub} . \square

To be able to apply Proposition 18, it remains to show that \succ^{sub} is effectively finitely branching. For this purpose, we first investigate the one-step relation \sqsubseteq_1^\emptyset induced by \sqsubseteq^\emptyset . Given an \mathcal{EL} concept C , we want to characterize the set of its upper neighbors $Upper(C) := \{D \mid C \sqsubseteq_1^\emptyset D\}$.

In a first step, we reduce the concept C by exhaustively replacing subconcepts of the form $E \sqcap F$ with $E \sqsubseteq^\emptyset F$ by E (modulo associativity and commutativity of \sqcap). As shown in (Küstners 2001), this can be done in polynomial time, and two concepts C, D are equivalent (i.e., $C \equiv^\emptyset D$) iff their reduced forms are equal up to associativity and commutativity of \sqcap .

Lemma 22. *Let C be a reduced \mathcal{EL} concept. Then, $Upper(C)$ can be computed in polynomial time. More precisely, it consists of the concepts D that can be obtained from C as follows:*

- Remove a concept name A from the top-level conjunction of C .
- Remove an existential restriction $\exists r.E$ from the top-level conjunction of C , and replace it by the conjunction of all existential restrictions $\exists r.F$ for $F \in Upper(E)$.

For example, if $C = A \sqcap \exists r.(B_1 \sqcap B_2 \sqcap B_3)$, then $Upper(C)$ consists of the two concepts $\exists r.(B_1 \sqcap B_2 \sqcap B_3)$ and $A \sqcap \exists r.(B_1 \sqcap B_2) \sqcap \exists r.(B_1 \sqcap B_3) \sqcap \exists r.(B_2 \sqcap B_3)$.

Unfortunately, this result does not transfer immediately from concept subsumption to axiom weakening. In fact, as we have seen before, strict subsumption need not produce a weaker axiom (see the remark below Proposition 20). Thus, to find all GCIs $C \sqsubseteq D'$ with $C \sqsubseteq D \succ_1^{sub} C \sqsubseteq D'$, it is not sufficient to consider only concepts D' with $D \sqsubseteq_1^\emptyset D'$. In case $C \sqsubseteq D'$ is equivalent to $C \sqsubseteq D$, we need to consider upper neighbors of D' , etc.

Proposition 23. *The one-step relation \succ_1^{sub} induced by \succ^{sub} is effectively finitely branching.*

Proof. Since \sqsubseteq^\emptyset is one-step generated, finitely branching, and well-founded, for a given concept D , there are only finitely many concepts D' such that $D \sqsubseteq^\emptyset D'$. Thus, a breadth first search along \sqsubseteq_1^\emptyset can be used to compute all concepts D' for which there is a path $D \sqsubseteq_1^\emptyset D_1 \sqsubseteq_1^\emptyset \dots \sqsubseteq_1^\emptyset D_n \sqsubseteq_1^\emptyset D'$ such that $C \sqsubseteq D$ is equivalent to $C \sqsubseteq D_i$ for $i = 1, \dots, n$, and $C \sqsubseteq D \succ^{sub} C \sqsubseteq D'$. Since \sqsubseteq^\emptyset is one-step generated, it is easy to see that all axioms γ with $C \sqsubseteq D \succ_1^{sub} \gamma$ can be obtained this way. However, the computed set of axioms may contain elements that are not one-step successors of $C \sqsubseteq D$. Thus, in a final step, we remove all axioms that are weaker than some axiom in the set. \square

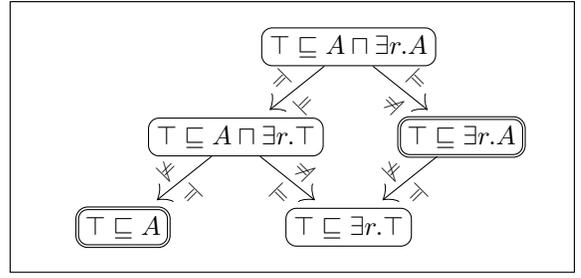


Figure 1: One-step weakening

Example 24. *To see that the final step of removing axioms in the last proof is really needed, consider the axiom $\beta = \top \sqsubseteq A \sqcap \exists r.A$ shown in Figure 1. The right-hand side $A \sqcap \exists r.A$ has two upper neighbors, namely $\exists r.A$ and $A \sqcap \exists r.T$. The first yields the axiom $\top \sqsubseteq \exists r.A$, which satisfies $\top \sqsubseteq A \sqcap \exists r.A \succ_1^{sub} \top \sqsubseteq \exists r.A$. The second yields the axiom $\top \sqsubseteq A \sqcap \exists r.T$, which is equivalent to β . The upper neighbor $\exists r.T$ of $A \sqcap \exists r.T$ yields the axiom $\top \sqsubseteq \exists r.T$, which is weaker than $\top \sqsubseteq \exists r.A$, and thus needs to be removed. In contrast, the upper neighbor A of $A \sqcap \exists r.T$ yields $\top \sqsubseteq A$, which satisfies $\top \sqsubseteq A \sqcap \exists r.A \succ_1^{sub} \top \sqsubseteq A$.*

A similar, but simpler example can be used to show that the additional removal of weaker elements in the proof of Proposition 18 is needed. Let α be the consequence $\top \sqsubseteq A$, $J = \{\beta\}$ for $\beta := \top \sqsubseteq A \sqcap B$, $\delta_1 := \top \sqsubseteq A$, $\delta_2 := \top \sqsubseteq B$, and $\gamma := \top \sqsubseteq \top$. Then we have exactly the situation described below the proof of Proposition 18, with \succ^{sub} as the employed weakening relation.

Corollary 25. *All maximally strong weakenings w.r.t. \succ^{sub} of an axiom in a justification can effectively be computed.*

Proof. By Proposition 18, this is an immediate consequence of the fact that \succ^{sub} is well-founded, one-step generated, and effectively finitely branching. \square

The algorithm for computing maximally strong weakenings described in the proof of Proposition 18 has non-elementary complexity for \succ^{sub} . In fact, the bound for the depth of the tree that must be searched grows by one exponential for every increase in the role-depth of the concept on the right-hand side. It is not clear how to obtain an algorithm with a better complexity. Example 30 below yields an exponential lower-bound, which still leaves a huge gap. We can also show that even deciding whether a given axiom is a maximally strong weakening w.r.t. \succ^{sub} is coNP-hard.

Syntactic Generalization

In order to obtain a weakening relation that has better algorithmic properties than \succ^{sub} , we consider a syntactic approach for generalizing \mathcal{EL} concepts. Basically, the concept D is a syntactic generalization of the concept C if D can be obtained from C by removing occurrences of subconcepts. To ensure that such a removal really generalizes the concept, we work here with reduced concepts.

Definition 26. *Let C, D be \mathcal{EL} concepts. Then D is a syntactic generalization of C (written $C \sqsubseteq^{syn} D$) if it is obtained*

from the reduced form of C by replacing some occurrences of subconcepts $\neq \top$ with \top .

For example, the concept $C = A_1 \sqcap \exists r.(A_1 \sqcap A_2)$ is already in reduced form, and its syntactic generalizations include $\top \sqcap \exists r.(A_1 \sqcap A_2) \equiv^{\emptyset} \exists r.(A_1 \sqcap A_2)$, $A_1 \sqcap \exists r.(\top \sqcap A_2) \equiv^{\emptyset} A_1 \sqcap \exists r.A_2$, $\exists r.\top$, and \top .

Lemma 27. *If $C \sqsubseteq^{syn} D$, then $C \sqsubseteq^{\emptyset} D$, and the length of any \sqsubseteq^{syn} -chain issuing from C is linearly bounded by the size of C .*

Proof. Since the concept constructors of \mathcal{EL} are monotonic, $C \sqsubseteq^{syn} D$ implies $C \sqsubseteq^{\emptyset} D$. Both the linear bound on the length of \sqsubseteq^{syn} -chains and strict subsumption follow from the fact that the m-size of the reduced form of C is strictly larger than the m-size of the reduced form of D , where the m-size counts only occurrences of concept and role names (see (Baader et al. 2018) for details). \square

By Proposition 15, this linear bound implies that \sqsubseteq^{syn} is one-step generated. A step in the corresponding one-step relation \sqsubseteq_1^{syn} can be realized by replacing a single concept name A or $\exists r.\top$ with \top . Though not all such single replacements are really in \sqsubseteq_1^{syn} , this is sufficient to show that \sqsubseteq^{syn} is effectively linearly branching.

Now, we define our new weakening relation, which syntactically generalizes the right-hand sides of GCIs:

$$C \sqsubseteq D \succ^{syn} C' \sqsubseteq D' \text{ if } C = C', D \sqsubseteq^{syn} D' \text{ and } \{C' \sqsubseteq D'\} \not\sqsubseteq C \sqsubseteq D.$$

The following theorem is an easy consequence of the properties of \sqsubseteq^{syn} and of our characterization of \sqsubseteq_1^{syn} .

Theorem 28. *The relation \succ^{syn} on \mathcal{EL} axioms is a linear, complete, one-step generated, and effectively linearly branching weakening relation.*

Due to the fact that \succ_1^{syn} -steps do not increase the size of axioms, the linear bounds on the branching of \succ_1^{syn} and the length of \succ^{syn} -chains imply that the algorithm described in the proof of Proposition 18 has an exponential search space.

Corollary 29. *All maximally strong weakenings w.r.t. \succ^{syn} can be computed in exponential time.*

The following example shows that there may be exponentially many maximally strong weakenings w.r.t. \succ^{syn} , and thus the exponential complexity stated above is optimal.

Example 30. *Let $\beta_i := P_i \sqcap Q_i \sqsubseteq B$ for $i = 1, \dots, n$ and $\beta := A \sqsubseteq P_1 \sqcap Q_1 \sqcap \dots \sqcap P_n \sqcap Q_n$. Consider $\mathfrak{D} = \mathfrak{D}_s \cup \mathfrak{D}_r$, where $\mathfrak{D}_s := \{\beta_i \mid 1 \leq i \leq n\}$ and $\mathfrak{D}_r := \{\beta\}$. Then $J = \{\beta\}$ is a justification for the consequence $\alpha = A \sqsubseteq B$, and all axioms of the form $A \sqsubseteq X_1 \sqcap X_2 \sqcap \dots \sqcap X_n$ ($X_i \in \{P_i, Q_i\}$) are maximally strong weakenings w.r.t. \succ^{syn} of β in J . The same is true for \succ^{sub} since in the absence of roles, these two weakening relations coincide.*

In contrast, computing a single maximally strong weakening is tractable.

Proposition 31. *A single maximally strong weakening w.r.t. \succ^{syn} can be computed in polynomial time.*

Proof. The algorithm that computes a maximally strong weakening works as follows. Starting from the concept $D' := \top$, it looks at all possible ways of making one step in the direction of D using \sqsubseteq_1^{syn} , i.e., it considers all D'' where $D \sqsubseteq^{syn} D'' \sqsubseteq_1^{syn} D'$. The concepts D'' can be obtained by adding a concept name A or an existential restriction $\exists r.\top$ at a place where (the reduced form of) D has such a concept or restriction. Obviously, there are only polynomially many such concepts D'' . For each of them we check whether $\mathfrak{D}_s \cup (J \setminus \{C \sqsubseteq D\}) \cup \{C \sqsubseteq D''\} \models \alpha$. If this is the case for all D'' , we return $C \sqsubseteq D'$. Otherwise, we choose an arbitrary D'' with $\mathfrak{D}_s \cup (J \setminus \{C \sqsubseteq D\}) \cup \{C \sqsubseteq D''\} \not\models \alpha$, and continue with $D' := D''$.

This algorithm terminates after linearly many iterations since in each iteration the size of D' is increased and it cannot get larger than D . In addition, $C \sqsubseteq D'$ is maximally strong since for each axiom $C \sqsubseteq E$ with $C \sqsubseteq D \succ^{syn} C \sqsubseteq E \succ^{syn} C \sqsubseteq D'$ there is a sequence $E \sqsubseteq_1^{syn} \dots \sqsubseteq_1^{syn} D'' \sqsubseteq_1^{syn} D'$. Consequently, $C \sqsubseteq D''$ has the consequence, and thus also $C \sqsubseteq E$. \square

Nevertheless, we can show that deciding whether an axiom is a maximally strong weakening w.r.t. \succ^{syn} is coNP-complete.

Conclusions

We have introduced a framework for repairing DL-based ontologies that is based on weakening axioms rather than deleting them, and have shown how to instantiate this framework for the DL \mathcal{EL} using appropriate weakening relations. Computing maximally strong weakenings w.r.t. these relations using the algorithm described in the proof of Proposition 18 is analogous to the black-box approach for computing justifications. It would be interesting to see whether a glass-box approach that modifies an \mathcal{EL} reasoning procedure can also be used for this purpose, similar to the way a tableau-based algorithm for \mathcal{ALC} was modified in (Lam et al. 2008).

Our weakening relations can also be used in the setting where the ontology is first modified, and then repaired using the classical approach as in (Du, Qi, and Fu 2014). In fact, for effectively finitely branching and well-founded weakening relations such as \succ^{sub} and \succ^{syn} , we can add for each axiom all (or some of) its finitely many weakenings w.r.t. the given relation, and then apply the classical repair approach. In contrast to the gentle repair approach proposed in this paper, a single axiom could then be replaced by several axioms, which might blow up the size of the ontology.

The standard reasoning procedures for \mathcal{EL} first normalize the given TBox, where normalization breaks up large GCIs into smaller ones (Baader et al. 2017). In some cases, applying classical repair to the normalized TBox also leads to more gentle repairs. For example, consider the refutable TBox $\mathcal{T} = \{A \sqsubseteq B_1 \sqcap B_2\}$, the strict ABox $\mathcal{A} = \{A(a)\}$, and the consequence $\alpha = (B_1 \sqcap B_2)(a)$. The TBox \mathcal{T} is normalized to $\mathcal{T}' = \{A \sqsubseteq B_1, A \sqsubseteq B_2\}$, which has the two classical repairs $\mathcal{T}'_1 = \{A \sqsubseteq B_1\}$ and $\mathcal{T}'_2 = \{A \sqsubseteq B_2\}$. This is exactly what our gentle repair approach (using \succ^{sub} or \succ^{syn}) would yield. However, normalization does not always do the job as illustrated by the following two examples. As a

first example, consider the refutable TBox $\{A \sqsubseteq \exists r.B\}$, the strict ABox $\{A(a)\}$, and the consequence $\exists r.B(a)$. Here, the TBox is normalized, and classical repair removes the GCI. In contrast, our gentle repair approach can weaken the GCI to $A \sqsubseteq \exists r.\top$. Another problem with using normalization in this setting is that in general it introduces new concept names. As a second example, consider the refutable TBox $\{A \sqsubseteq \exists r.\exists r.B\}$ and the strict ABox $\{A(a)\}$, where the unwanted consequence is $\exists r.\exists r.B(a)$. Normalizing the TBox yields $\{A \sqsubseteq \exists r.X, X \sqsubseteq \exists r.B\}$; thus, classical repair yields as repairs the TBoxes consisting of $A \sqsubseteq \exists r.X$ or $X \sqsubseteq \exists r.B$. These two axioms do not make sense for the user since X is a name without meaning in the application. Thus, some post-processing that can get rid of the new names (similar to forgetting (Nikitina and Rudolph 2014)) would be required. While an approach based on appropriate variants of normalization and forgetting may be able to generate gentle repairs akin to what our approach produces using \succ^{syn} , it would not be able to deal with more sophisticated weakening relations such as \succ^{sub} . In addition, classical repair applied to the normalized TBox would not distinguish between more or less gentle repairs, and would also produce all classical repairs of the original TBox.

Acknowledgments

This work is partially supported by DFG within the Research Training Group 1907 (RoSI). The authors would like to thank Bernhard Ganter for helpful discussion regarding one-step generated pre-orders.

References

- Baader, F., and Hollunder, B. 1995. Embedding defaults into terminological knowledge representation formalisms. *J. of Autom. Reas.* 14:149–180.
- Baader, F., and Morawska, B. 2010. Unification in the description logic \mathcal{EL} . *Logical Methods in Computer Science* 6(3).
- Baader, F., and Suntisrivaraporn, B. 2008. Debugging SNOMED CT using axiom pinpointing in the description logic \mathcal{EL}^+ . In *Proc. KR-MED 2008*.
- Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.
- Baader, F.; Kriegel, F.; Nuradiansyah, A.; and Peñaloza, R. 2018. Repairing description logic ontologies by weakening axioms. LTCS-Report 18-01, TU Dresden. See <http://lat.inf.tu-dresden.de/research/reports.html>.
- Baader, F.; Borchmann, D.; and Nuradiansyah, A. 2017. The identity problem in description logic ontologies and its application to view-based information hiding. In *Proc. JIST 2017*, volume 10675 of *LNCS*, 102–117. Springer-Verlag.
- Baader, F.; Peñaloza, R.; and Suntisrivaraporn, B. 2007. Pinpointing in the description logic \mathcal{EL}^+ . In *Proc. KI 2007*, volume 4667 of *LNAI*, 52–67. Springer-Verlag.
- Brandt, S. 2004. Polynomial time reasoning in a description logic with existential restrictions, GCI axioms, and—what else? In *Proc. ECAI 2004*, 298–302.
- Cuenca Grau, B. 2010. Privacy in ontology-based information systems: A pending matter. *Semantic Web* 1:137–141.
- Du, J.; Qi, G.; and Fu, X. 2014. A practical fine-grained approach to resolving incoherent OWL 2 DL terminologies. In *Proc. CIKM 2014*, 919–928.
- Garey, M. R., and Johnson, D. S. 1979. *Computers and Intractability — A guide to NP-completeness*. W. H. Freeman and Company.
- Horridge, M.; Parsia, B.; and Sattler, U. 2008. Laconic and precise justifications in OWL. In *Proc. ISWC 2008*, volume 5318 of *LNCS*, 323–338. Springer-Verlag.
- Kalyanpur, A.; Parsia, B.; Horridge, M.; and Sirin, E. 2007. Finding all justifications of OWL DL entailments. In *Proc. ISWC 2007*, volume 4825 of *LNCS*, 267–280. Springer-Verlag.
- Küsters, R. 2001. *Non-standard Inferences in Description Logics*, volume 2100 of *LNAI*. Springer-Verlag.
- Lam, J. S. C.; Sleeman, D.; Pan, J. Z.; and Vasconcelos, W. 2008. A fine-grained approach to resolving unsatisfiable ontologies. In *Journal on Data Semantics X*, volume 4900 of *LNCS*, 62–95. Springer-Verlag.
- Lehmann, J., and Hitzler, P. 2010. Concept learning in description logics using refinement operators. *Machine Learning* 78(1-2):203–250.
- Meyer, T.; Lee, K.; Booth, R.; and Pan, J. Z. 2006. Finding maximally satisfiable terminologies for the description logic \mathcal{ALC} . In *Proc. AAI 2006*. AAAI Press/The MIT Press.
- Nikitina, N., and Rudolph, S. 2014. (Non-)succinctness of uniform interpolants of general terminologies in the description logic \mathcal{EL} . *Artif. Intell.* 215:120–140.
- Parsia, B.; Sirin, E.; and Kalyanpur, A. 2005. Debugging OWL ontologies. In *Proc. WWW 2005*, 633–640. ACM.
- Reiter, R. 1987. A theory of diagnosis from first principles. *Artif. Intell.* 32(1):57–95.
- Schlobach, S., and Cornet, R. 2003. Non-standard reasoning services for the debugging of description logic terminologies. In *Proc. IJCAI 2003*, 355–362. Morgan Kaufmann, Los Altos.
- Schlobach, S.; Huang, Z.; Cornet, R.; and Harmelen, F. 2007. Debugging incoherent terminologies. *J. Automated Reasoning* 39(3):317–349.
- Troquard, N.; Confalonieri, R.; Galliani, P.; Peñaloza, R.; Porello, D.; and Kutz, O. 2018. Repairing ontologies via axiom weakening. In *Proc. AAI 2018*. AAAI Press.