

Ontological Analysis For Description Logics Knowledge Base Debugging

Julien Corman
IRIT, Toulouse, France

Nathalie Aussenac-Gilles
CNRS - IRIT, Toulouse, France

Laure Vieu
CNRS - IRIT, Toulouse, France
LOA, Trento, Italy

Abstract

Formal ontology provides axiomatizations of domain independent principles which, among other applications, can be used to identify modeling errors within a knowledge base. The Ontoclean methodology is probably the best-known illustration of this strategy, but its cost in terms of manual work is often considered dissuasive. This article investigates the applicability of such debugging strategies to Description Logics knowledge bases, showing that even a partial and shallow analysis rapidly performed with a top-level ontology can reveal the presence of violations of common sense, and that the bottleneck, if there is one, may instead reside in the resolution of the resulting inconsistency or incoherence.

Important efforts have been devoted in the Description Logics (DL) community to the development of tools or algorithms to manipulate and share knowledge expressed as sets of logical statements. One of the obstacles to the adoption of these technologies is the reliability of such knowledge bases (KB) as they grow in size. Although each statement within a KB may make sense individually, several of them can easily lead to undesired inferences, even if the KB is logically consistent and coherent,¹ due to incompatible understandings of elements of its signature. Identifying and repairing such errors is of particular importance for semantic web applications, which rely upon the reuse of data possibly aggregated from different sources, as well as (modules of) often weakly axiomatized data schemas. As an illustration, consider the following consistent set of DL statements, extracted from DBPedia (Mendes, Jakob, and Bizer 2012) :

Ex 1.

$K = \{(1) \text{ keyPerson}(\textit{Virgin Holidays}, \textit{CEO}),$
 $(2) \text{ keyPerson}(\textit{Caixa Bank}, \textit{CEO}),$
 $(3) \text{ occupation}(\textit{Peter Munk}, \textit{CEO}),$
 $(4) \top \sqsubseteq \forall \text{keyPerson}. \text{Person},$
 $(5) \text{ keyPerson}(\textit{BrookField Office Properties},$
 $\textit{Peter Munk})\}$

It is relatively easy to find a plausible interpretation of statements 1, 2 and 3 taken together. But adding either 4 or 5 immediately violates simple intuitions. Some additional

knowledge Δ is needed though in order to spot it on a formal basis. For instance, if $\Delta \models \top \sqsubseteq \forall \text{occupation}. \neg \text{Person}$, (“nothing can be both a person and the occupation of someone”), then both $\{1,3,4\}$ and $\{2,3,4\}$ would contradict it.

Formal ontology, as a discipline, provides axiomatizations of domain independent principles which can be used as the core of this additional knowledge Δ . Let for instance Ω be a foundational ontology, and $\Omega \models \text{Function} \sqsubseteq \neg \text{PhysicalEntity}$. In order to obtain a set $\Delta \models \top \sqsubseteq \forall \text{occupation}. \neg \text{Person}$, Ω can be extended with an “anchoring” set Γ of formulas that relates elements of the signature of Ω to elements of the signature of the input KB K , such that $\Delta = \Omega \cup \Gamma$. A possible Γ here would be $\Gamma = \{\text{Person} \sqsubseteq \text{PhysicalEntity}, \top \sqsubseteq \forall \text{occupation}. \text{Function}\}$.

The best-known example of this approach is probably the Ontoclean methodology (Guarino and Welty 2000), which provides a set Ω of constraints over the subsumption relation in a taxonomy, based on properties of its atomic concepts, expressed as second-order predicates : (non-)identity, (anti-/non-)rigidity, ... Γ in this case is an assignment of these properties to concepts of the input KB. But manually crafting such a Γ has been pointed out as costly and error-prone, with a possibly low inter-annotator agreement, as illustrated in (Völker et al. 2008).

This article investigates the usage of a first-order foundational ontology as Ω , and shows that the bottleneck, if there is one, may not be the crafting of Γ , which can be reduced to a few simple and limited judgments avoiding complex ontological questions, but instead the resolution of the resulting inconsistency/incoherence due to coexisting incompatible conceptualizations. This is supported by the evaluation of a partial and shallow ontological analysis in section 2, followed by a critical review of existing revision operators for DLs in section 3.

1 Preliminaries

The reader is assumed familiar with the syntax and standard model-theoretic semantic of DLs, as well as the DL based languages of the OWL family. For an introduction to DLs, see (Baader 2003). A KB is understood as a finite set of DL formulas, also called *axioms*.

The signature $sig(X)$ of a set X of DL statements is defined as $sig(X) = \mathcal{A}_X \cup \mathcal{E}_X \cup \mathcal{R}_X$, with \mathcal{A}_X , \mathcal{E}_X and \mathcal{R}_X mutually disjoint, standing respectively for the sets of DL

atomic concepts, DL individuals and DL roles appearing in X . Inconsistency is understood in the traditional way, i.e. X is inconsistent iff it has no model. Another commonly required property of a DL KB is *coherence*. An atomic concept $A \in \mathcal{A}_X$ is *unsatisfiable* wrt X iff the interpretation of A is empty in all models of X . Then X is said *incoherent* iff there is an $A \in \mathcal{A}_K$ unsatisfiable wrt X .

OWL *datatype properties* are treated as (fresh) DL atomic concepts : for instance, the assertion “hasDensity(Iowa, 54.8)” is understood as the ABox statement “hasDensity(Iowa)”, regardless of the value.

2 Ontological analysis

2.1 Using a top-level ontology as a debugging tool

This section focuses on the choice of the core ontological knowledge Ω , and on the constitution of the set Γ of anchoring formulas. The Ontoclean methodology as implemented by (Glimm, Rudolph, and Völker 2010) or (Welty 2006) provides a set Ω of constraints on the subsumption relation, expressed wrt (second-order) properties of the involved concepts, for instance “a rigid concept cannot be subsumed by an anti-rigid concept”. Therefore Γ must entail predications over reified concepts, like “Rigid(*concept*₁)” or “AntiRigid(*concept*₂)”.

An arguably simpler approach is investigated here, which consists in tagging an automatically selected subset E of $\text{sig}(K)$, using as the core ontological knowledge Ω a (DL expressible subset of a) top-level ontology, like DOLCE or BFO (both described in (Masolo et al. 2002)), which provide axiomatizations of fundamental first-order categories, like Event, PhysicalEntity, Function, ... Γ here is composed of manually crafted statements relating elements of $\text{sig}(K)$ to elements of $\text{sig}(\Omega)$, of the form $B(x)$, $A \sqsubseteq B$, $\top \sqsubseteq \forall R.B$, or $\exists R. \top \sqsubseteq B$, with $x \in \mathcal{E}_K$, $A \in \mathcal{A}_K$, $R \in \mathcal{R}_K$, and B (one of) the most specific element(s) of \mathcal{A}_Ω which intuitively verify(ies) the statement. This analysis is not meant to reproduce another one performed with Ontoclean though, as both may spot different errors. The scope in particular is larger in that the whole KB (ABox + TBox) can be investigated, and not exclusively the taxonomy.

An advantage of choosing such an Ω is that Γ does not quantify over reified concepts or relations, but consists instead of simple DL formulas quantifying over the domain modeled by K , like “Person \sqsubseteq PhysicalEntity”, “Function(*CEO*)” or “ $\top \sqsubseteq \forall \text{hasKeyPerson. PhysicalEntity}$ ”. Predicating and quantifying over the individuals of the intended domain is arguably easier than assigning second order properties to concepts (or relations).

Additionally, in order to reduce the amount of required manual work, we experimented the analysis of only a small subset E of $\text{sig}(K)$, based on frequency within K , the underlying assumption being that an element of $\text{sig}(K)$ is more likely to be used in intuitively incompatible ways if it appears in more axioms.

2.2 Evaluation

The core ontological knowledge Ω for the evaluation was the ontology backing the categorization tool TME0 used in the

Senso Comune project (Jezek et al. 2014). It is a simplified and slightly extended variant of the top-level DOLCE, formalized as DL expressible first-order constraints. The main input DL KB was automatically retrieved from a DBpedia endpoint² in October 2013. We first tried to extract a thematically coherent ABox, taking advantage of Wikipedia categories for the retrieval,³ and favoring overall connectedness between individuals.⁴ All subsuming classes, as well as domains and ranges, were then retrieved recursively, yielding the KB K_{DBP} , composed of 6867 logical axioms, with 106 atomic concepts, 1361 individuals, 100 objectProperties and 205 dataProperties, the least expressive underlying DL being $\mathcal{AL}^{(D)}$.

In order to select the set E of elements of $\text{sig}(K)$ which were candidates for appearing in Γ , each of \mathcal{A}_K , \mathcal{E}_K and \mathcal{R}_K was ordered by decreasing number of syntactic occurrences within K . This heuristic was refined by quotienting K before computing these occurrences, taking into account the diversity of the axioms each element appears in, such that for instance the axioms “location(*Wang Theatre, Boston*)” and “location(*Ether Dome, Boston*)” count as one occurrence only for the element *Boston*. An integer parameter n was used to test different sizes for Γ , such that E contained the n most frequent elements of each of \mathcal{A}_K , \mathcal{E}_K , and \mathcal{R}_K . Setting n to 2, with only 5 axioms in Γ , was sufficient to obtain an inconsistent $K \cup \Omega \cup \Gamma$.

In order evaluate recall as well, two gold standards were build out of two automatically generated subsets K_1 and K_2 of K_{DBP} , of 104 and 96 axioms respectively, each centered on one individual with high frequency in K_{DBP} . Within each K_i , a set $\sigma(K_i)$ of erroneous axioms was identified by a formal ontology expert (11 axioms for K_1 , and 7 for K_2). An axiom ϕ was considered erroneous iff the understanding of some element of $\text{sig}(\{\phi\})$ was incompatible with its overall understanding within K_{DBP} . For instance, the axiom “director(*Museum Of The Rockies, Smithsonian Institution*)” was considered erroneous because, majoritarilly within K_{DBP} , the individual *Smithsonian Institution* is understood as an organization, whereas the DL role director is understood as ranging over human beings. The evaluation consisted in checking whether these erroneous axioms were actually involved in the inconsistency/incoherence of $K_i \cup \Omega \cup \Gamma$. For each K_i , the library developed by (Horridge 2011) was used to compute the set \mathcal{J} of all justifications for the inconsistency/incoherence, i.e. all set-inclusion minimal inconsistent/incoherent sets of axioms of $K_i \cup \Omega \cup \Gamma$. For K_1 , setting $n = 6$ (with $|\Gamma| = 12$) was sufficient to obtain $|\sigma(K_1) \cap \bigcup \mathcal{J}| = 7$, i.e. to identify 7 of the 11 erroneous axioms as potentially responsible for the inconsistency/incoherence. For K_2 , setting $n = 11$ ($|\Gamma| = 29$) was sufficient to cover all 7 erroneous axioms.

²<http://dbpedia-live.openlinksw.com/sparql>

³but they were not incorporated into the KB

⁴An additional precaution was taken for the ABox. The dataset available from the endpoint has been procedurally closed under the rule $\{A(x), A \sqsubseteq B\} \vdash_d B(x)$, with A and B atomic concepts. In order to be closer to the original syntactic formulation, each formula ϕ such that $K \setminus \phi \vdash_d \phi$ holds was removed from K .

For each of K_{DBP} , K_1 , K_2 , the constitution of Γ with the above values for n took less than an hour, which shows that a relatively efficient ontological analysis can be performed in this case with a limited amount of manual work.

3 Repair

This section focuses on the automated suggestion of weaker versions of $K \cup \Gamma$, in order to accommodate for the core ontological knowledge Ω (possible extensions of K after this weakening phase are beyond the scope of this article). By default, it will be assumed that the anchoring set Γ is not more reliable than K , but nothing prevents considering part of it as safe, in which case Ω can be extended with it.

An intuitive requirement is that no information should be unnecessarily lost in the process. This problem has been extensively studied for propositional logic in the field of belief change, and different proposals have been made more recently to adapt belief change operators and postulates to DLs. Due to the lack of place, the discussion will remain informal, excluding limit cases on purpose (e.g. Ω is assumed to be consistent). For a comprehensive overview of belief change postulates and representation theorems for DLs, see (Ribeiro 2013).

The process can be viewed as the result of a *revision*⁵ of $K \cup \Gamma$ by Ω , written $(K \cup \Gamma) \star \Omega$, followed by the isolation of what was initially implied by K , i.e. $Cn(K) \cap ((K \cup \Gamma) \star \Omega)$.

Limitations for this task of the various principles of minimal information loss proposed for DLs are illustrated in what follows.

3.1 Semantic notions of minimal information loss

The dominant scheme in the belief change literature considers *belief sets*, i.e. (deductively closed) theories, or equivalently sets of models, and a common way of defining minimal information loss in semantic terms relies on some distance between interpretations. In this view, $(K \cup \Gamma) \star \Omega$ is defined as the belief set whose models verify Ω and are the closest to models of $K \cup \Gamma$. But adopting a distance between interpretations for DLs is less straightforward than for propositional logic. In example 2, all other things being equal, an interpretation \mathcal{I}_1 verifying the ABox statement “ $\exists \text{training}. \top (\text{Mabel Esplin})$ ” seems intuitively closer to the models of K than an interpretation \mathcal{I}_2 verifying “ $\forall \text{training}. \perp (\text{Mabel Esplin})$ ”. But to our knowledge, none of the distances which have been proposed for revision in DLs favors \mathcal{I}_1 over \mathcal{I}_2 (neither the distance used in (Qi and Du 2009), nor the syntactico-semantic distance used by (Meyer, Lee, and Booth 2005; Qi, Liu, and Bell 2006)).

Ex 2.

$$\begin{aligned}
 K &= \{ \text{training}(\text{Mabel Esplin}, \text{Christopher Whall}), \\
 &\quad \top \sqsubseteq \forall \text{training}. \text{EducationalInstitution} \} \\
 \Omega &\models \text{PhysicalEntity} \sqsubseteq \neg \text{NonPhysicalEntity} \\
 \Gamma &= \{ \text{EducationalInstitution} \sqsubseteq \text{NonPhysicalEntity}, \\
 &\quad \text{PhysicalEntity}(\text{Christopher Whall}) \}
 \end{aligned}$$

⁵The operation of belief *contraction* is arguably more intuitive here. But it requires some form of negation of Ω , which is not straightforward in DLs (Flouris et al. 2006).

3.2 Syntactic notions of minimal information loss

A more common approach to DL KB debugging focuses on belief bases, i.e. sets of axioms. In this view, if \mathcal{M} is the set of set-inclusion maximal subbases of $K \cup \Gamma$ consistent/coherent with Ω , minimizing information loss has been defined as selecting either one element of \mathcal{M} , or the intersection or disjunction (Meyer, Lee, and Booth 2005) of several of them. Different algorithms have been designed for DL belief base revision (or contraction) (Schlobach 2005; Qi et al. 2008; Ribeiro and Wassermann 2008; Friedrich and Shchekotykhin 2005; Kalyanpur et al. 2006), most of them based on diagnosis (Reiter 1987). As experimented by (Schlobach 2005), a practical limitation of syntax dependent KB debugging is the cardinality of \mathcal{M} . Selecting one of its elements arbitrarily is not a viable option, whereas taking the conjunction, or even the disjunction of all of them tends to yield a KB that is too weak. We verified this observation by computing a lower bound on $|\mathcal{M}|$ with a slightly modified version of Reiter’s algorithm. For K_{DBP} , with $n = 6$, we obtained $|\mathcal{M}| \geq 174$. Preserving the anchoring set Γ did not solve the problem, with $|\mathcal{M}| \geq 180$. Even for the small KBs K_1 and K_2 , with $n = 6$ and $n = 11$ respectively, we obtained $|\mathcal{M}| \geq 340$ and $|\mathcal{M}| \geq 270$, and $|\mathcal{M}| \geq 32$ and $|\mathcal{M}| \geq 36$ when preserving Γ .

To address this issue, (Qi et al. 2008) or (Kalyanpur et al. 2006) proposed to compute cardinality maximal (and not only set-inclusion maximal) subsets of $K \cup \Gamma$ consistent/coherent with Ω . On relatively large KBs like K_{DBP} though (> 6 000 axioms), it seems quite arbitrary to claim *a priori* that the removal of 10 axioms only should be preferred to the removal of 11 or even 15 of them. We also empirically verified for K_1 and K_2 that the maximal cardinality heuristic failed to discard the axioms of $\sigma(K_i)$, with $n = 6$ and $n = 11$ respectively, even when preserving Γ . To this end, we first checked that removing from K all erroneous axioms appearing in some justification was not sufficient to restore consistency. In such a case, an optimal candidate for removal is a set-inclusion minimal diagnosis that contains as many erroneous axioms as possible. For K_1 , one of the subbases yielded by the maximal cardinality heuristic was actually the complement of the only optimal candidate for removal. But there were 31 other candidate subbases. For K_2 , none of the optimal candidates was discarded. As an alternative, (Schlobach and Cornet 2003) or (Qi et al. 2008) used the notion of *core*, prioritizing the removal of axioms which appear in a higher number of justifications. Unfortunately, this is generally correlated with the cardinality heuristic just mentioned.

Belief base revision seems nonetheless a promising option for the automated weakening of $K \cup \Gamma$, provided an adequate method to select elements of \mathcal{M} , or to compute a preference relation over the axioms potentially involved in the inconsistency/incoherence of $K \cup \Gamma \cup \Omega$ (i.e. $\bigcup \mathcal{J}$). We are currently experimenting the usage of automatically gathered linguistic evidence for both.

References

- Baader, F. 2003. *The description logic handbook: theory, implementation, and applications*. Cambridge university press.
- Flouris, G.; Huang, Z.; Pan, J. Z.; Plexousakis, D.; and Wache, H. 2006. Inconsistencies, negations and changes in ontologies. In *NCAI*.
- Friedrich, G., and Shchekotykhin, K. 2005. A general diagnosis method for ontologies. In *The Semantic Web–ISWC 2005*. Springer. 232–246.
- Glimm, B.; Rudolph, S.; and Völker, J. 2010. Integrated metamodeling and diagnosis in OWL 2. In *The Semantic Web–ISWC 2010*. Springer.
- Guarino, N., and Welty, C. 2000. A formal ontology of properties. In *EKAW*.
- Horrige, M. 2011. *Justification based explanation in ontologies*. Ph.D. Dissertation, the University of Manchester.
- Jezek, E.; Vieu, L.; Zanzotto, F.; Vetere, G.; Oltramari, A.; Gangemi, A.; and Varvara, R. 2014. Extending ‘senso comune’ with semantic role sets. In Bunt, H., ed., *10th Workshop on Interoperable Semantic Annotation (ISA-10)*.
- Kalyanpur, A.; Parsia, B.; Sirin, E.; and Cuenca-Grau, B. 2006. Repairing unsatisfiable concepts in OWL ontologies. In *Proceedings of the 3rd European conference on The Semantic Web: research and applications*. Springer-Verlag.
- Masolo, C.; Borgo, S.; Gangemi, A.; Guarino, N.; Oltramari, A.; Oltramari, R.; Schneider, L.; Istituzioni, L. P.; and Horrocks, I. 2002. Wonderweb deliverable d17. the wonderweb library of foundational ontologies and the dolce ontology.
- Mendes, P. N.; Jakob, M.; and Bizer, C. 2012. DBpedia: A multilingual cross-domain knowledge base. In *LREC*.
- Meyer, T.; Lee, K.; and Booth, R. 2005. Knowledge integration for description logics. In *NCAI*.
- Qi, G., and Du, J. 2009. Model-based revision operators for terminologies in description logics. In *IJCAI*.
- Qi, G.; Haase, P.; Huang, Z.; Ji, Q.; Pan, J. Z.; and Völker, J. 2008. A kernel revision operator for terminologies - algorithms and evaluation. In *ISWC*.
- Qi, G.; Liu, W.; and Bell, D. 2006. A revision-based approach to handling inconsistency in description logics. *Artificial Intelligence Review* 26(1-2).
- Reiter, R. 1987. A theory of diagnosis from first principles. *Artificial intelligence* 32(1).
- Ribeiro, M. M., and Wassermann, R. 2008. The ontology reviser plug-in for protégé. In *WONTO*.
- Ribeiro, M. M. 2013. *Belief revision in non-classical logics*. Springer.
- Schlobach, S., and Cornet, R. 2003. Non-standard reasoning services for the debugging of description logic terminologies. In *IJCAI*.
- Schlobach, S. 2005. Diagnosing terminologies. In *AAAI*.
- Völker, J.; Vrandečić, D.; Sure, Y.; and Hotho, A. 2008. AEON—an approach to the automatic evaluation of ontologies. *Applied Ontology* 3(1).
- Welty, C. 2006. OntOWLclean: Cleaning OWL ontologies with OWL. *Frontiers in artificial intelligence and applications* 150.