How to Contract Ontologies

Statement of Interest

Bernardo Cuenca Grau[1], Evgeny Kharlamov[2], and Dmitriy Zheleznyakov[2]

[1] Department of Computer Science, University of Oxford
bernardo.cuenca.grau@cs.ox.ac.uk
[2] KRDB Research Centre, Free University of Bozen-Bolzano
{kharlamov,zheleznyakov}@inf.unibz.it

Abstract. The dynamic nature of knowledge development has motivated the formal study of ontology evolution problems. In this paper, we study ontology contraction—the problem of retracting information that is no longer considered to hold. Our prime interest is to focus on ontologies expressed in Description Logics *DL-Lite* and $\mathcal{EL}$, which underpin the OWL 2 QL and OWL 2 EL profiles. Our goal is to understand how to compute contractions under different kinds of semantics. As we have already shown in [1, 2], ontology contraction is technically very challenging.

1 Importance of Ontology Contraction

Ontologies written in the Web Ontology Language (OWL) [3] and its revision OWL 2 [4] are becoming increasingly important for a wide range of applications. The formal underpinning of OWL is based on Description Logics (DLs) – knowledge representation formalisms with well-understood computational properties [5]. A DL ontology $\mathcal{K}$ consists of a TBox $\mathcal{T}$, describing general (i.e., schema-level) domain knowledge, and an ABox $\mathcal{A}$, providing data about specific individuals.

Ontologies are not static entities, but rather they are frequently modified when new information needs to be incorporated, or existing information is no longer considered valid. The impact of such changes on the semantics of the ontology, however, is difficult to predict and understand. This dynamic nature of ontologies motivates the study of *ontology evolution problems* from both foundational and practical perspectives [2, 6–15]. We are interested in a particular aspect of ontology evolution, namely *contraction* – the process of "retracting" information that is no longer considered to hold [16, 17]. This information is typically represented by a formula α.

Two important forms of contraction are crucial for many ontology design and management tasks:

- *TBox contraction*, where the axiom α to be retracted is a TBox axiom; and
- *ABox contraction*, where α is an ABox assertion and the TBox of the original ontology should remain the same.

Consider scenarios for TBox and ABox contraction. OWL TBoxes are extensively used in the clinical sciences, and clinical ontologies such as SNOMED CT[3] and NCI[4] are subject to frequent modifications that involve retracting unintended consequences from the TBox [18]. For example, the developers of NCI perform over 900 monthly changes [18]. Furthermore, ABox contraction is important for applications relying on widely-used *reference TBoxes*. For example, bio-informaticians working on gene extraction can describe the experimental results using an ABox according to standard gene TBoxes. New experiments may imply that some facts about specific individuals no longer hold, which should be reflected in the ABox; at the same time, TBoxes should clearly not be affected by these manipulations of the data.

2 Practical View of Contraction

Approaches to ontology contraction, typically adopted in practice (especially when changes occur at the TBox-level), are essentially *syntactic* [9, 13, 19]. Many such approaches are based on the notion of a *justification*: a minimal subset of the ontology that entails a given consequence [20, 21]. For example, to retract an axiom α entailed by $\mathcal{K}$, it suffices to compute all justifications for α in $\mathcal{K}$, find a minimal subset $\mathcal{R}$ of $\mathcal{K}$ (a "repair") with at least one axiom from each justification, and take $\mathcal{K}_o = \mathcal{K} \setminus \mathcal{R}$ as the result of the contraction. Retracting α results in the deletion of a minimal set of axioms and hence the structure of $\mathcal{K}$ is maximally preserved.

3 Logic-Based View of Contraction

From a logic-based perspective, the desirable properties of contraction should be dictated by the *principle of minimal change* [16], according to which the semantics of the ontology should change "as little as possible", thus ensuring that the contraction has the least possible impact.

Logic-based semantics derived from the principle of minimal change have been recently studied in the more general context of ontology evolution. These semantics are either *model-based* (MBS) or *formula-based* (FBS). Under both types of semantics, evolution of an ontology $\mathcal{K}$ written in a DL $\mathcal{DL}$ results in a $\mathcal{DL}$-ontology $\mathcal{K}_o$ in which the required information is incorporated, retracted, or updated; the difference is in the way $\mathcal{K}_o$ is obtained. Under MBS the set of all models $\mathcal{M}$ of $\mathcal{K}$ evolves into a new set $\mathcal{M}'$ of models that are "as close as possible" to those in $\mathcal{M}$ (w.r.t. some notion of distance between models); then, $\mathcal{K}_o$ is the ontology that axiomatises $\mathcal{M}'$ [10–12, 22, 23]. Under FBS, $\mathcal{K}_o$ is an ontology defined in terms of the *deductive closure* of $\mathcal{K}$ that satisfies the evolution requirements. FBS, however, have been less studied in the context of ontologies [1, 2, 11, 24].

[3] http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html
[4] National Cancer Institute Thesaurus

Under both MBS and FBS, we are interested in computing an "optimal" contraction —that is, a contraction that is as similar as possible to the original ontology. In particular, an optimal contraction of a $\mathcal{DL}$-ontology $\mathcal{K}$ with an axiom α is a $\mathcal{DL}$-ontology $\mathcal{K}_o$ such that *(i)* $\mathcal{K}$ entails $\mathcal{K}_o$, *(ii)* $\mathcal{K}_o$ does not entail α, and *(iii)* $\mathcal{K}_o$ is "as similar as possible" to $\mathcal{K}$ according to the particular notion of minimal change adopted by the semantics under consideration.

4 Issues with Practical and Logic-Based Approaches

Practical approaches suffer from intrinsic information loss. More precisely, by removing $\mathcal{R}$ from $\mathcal{K}$, one may inadvertently retract consequences of $\mathcal{K}$ other than α, which are "intended". Identifying and recovering such intended consequences is an important issue. For example, if $\mathcal{K}$ consists of two TBox axioms: β_1 = "VW is-a car", β_2 = "car is-a vehicle", with an implicit information β_3 = "VW is-a vehicle", and $\alpha = \beta_1$, then $\mathcal{K}_o = \{\beta_2\}$; thus, β_3 is lost.

Logic-based approaches limit information loss, but they suffer from the following two problems. First, given a $\mathcal{DL}$-ontology $\mathcal{K}$ and an axiom α, an optimal contraction $\mathcal{K}_o$ may not exist in $\mathcal{DL}$, that is $\mathcal{DL}$ may not be *closed* under contraction. It has been shown [11, 22] that *DL-Lite* is not closed under so-called update and revision on both TBox and ABox levels; in [1] these results were extended to TBox and ABox contraction. In [1, 2] it has been shown that $\mathcal{EL}$ is not closed under TBox and ABox contraction.

The second problem is that logic-based approaches are also problematic from a modeling point of view. Indeed, in contrast to syntactic approaches, these semantics do not distinguish between the axioms in the deductive closure that are explicit in $\mathcal{K}$, and those that are merely implied. Ontologies, however, are the result of a time-consuming modeling process, and thus contractions should also preserve as much as possible the structure of $\mathcal{K}$. For example, consider a *DL-Lite*-ontology $\mathcal{K}$ with the following two axioms β_2 = "car is-a vehicle" and β_4 = "VW is-a car and golf", and let α = "lecturer is-a professor". Clearly α is unrelated to $\mathcal{K}$, and the optimal contraction of $\mathcal{K}$ with α is expected to be $\mathcal{K}$ itself. At the same time, the ontology $\mathcal{K}_o$ consisting of the following axioms is a valid contraction too (under any FBS and MBS considered in the ontology evolution literature): β_2 = "car is-a vehicle" and β_5 = "VW is-a car", β_6 = "VW is-a golf", β_7 = "VW is-a vehicle". Returning such $\mathcal{K}_o$ as the contraction result might be undesirable from a practical point of view.

We believe that our results in [1] suggest that classical approaches to ontology contraction, which are well-understood and well-behaved for propositional logics, are intrinsically problematical in the context of ontology languages.

5 Bridging Logics and Practice

In [2], these limitations of both syntactic and logic-based approaches were addressed (at least partly). On the one hand, the semantics in [2] provides a "bridge" between syntactic and formula-based approaches; on the other hand, this semantics provides a distinction between the languages $\mathcal{DL}$ in which the original ontology $\mathcal{K}$ and the resulting contraction are expressed, and the language $\mathcal{LP}$ (the *preservation language*), which expresses the entailments of $\mathcal{K}$ that must be maximally preserved. The principle of minimal change is reflected along two dimensions:

(i) *structural*, where the explicit axioms in $\mathcal{K}$ are maximally preserved;
(ii) *deductive*, where the consequences of $\mathcal{K}$ in $\mathcal{LP}$ are maximally preserved.

More formally, the semantics of [2] can be defined as follows in the context of contraction.

Definition 1. *Let $\mathcal{DL}$ and $\mathcal{LP}$ be description logics with $\mathcal{LP} \subseteq \mathcal{DL}$. Let $\mathcal{K}$ be a $\mathcal{DL}$-ontology, and let α be a $\mathcal{DL}$-axiom such that $\mathcal{K} \models \alpha$.*

A $\mathcal{DL}$-ontology $\mathcal{K}_o$ is an optimal contraction of $\mathcal{K}$ with α w.r.t. the preservation language $\mathcal{LP}$ if (i) $\mathcal{K} \models \mathcal{K}_o$; (ii) $\mathcal{K}_o \not\models \alpha$; (iii) $\mathcal{K}_o \cup \{\beta\} \models \alpha$, for each $\beta \in \mathcal{K} \setminus \mathcal{K}_o$; and (iv) $\mathcal{K}_o \cup \{\gamma\} \models \alpha$, for each $\mathcal{LP}$-axiom γ that is entailed by $\mathcal{K}$ but not by $\mathcal{K}_o$.

6 Our Goals

The contraction algorithms we proposed and implemented in [2] are specific to TBox contraction. We would like to extend them to cover more expressive DLs. Another next step is to develop algorithms for ABox contraction.

Acknowledgements. B. Cuenca Grau is supported by a Royal Society Fellowship. E. Kharlamov is supported by the EPSRC EP/G004021/1, EP/H017690/1, and ERC FP7 grant Webdam (n. 226513). E. Kharlamov and D. Zheleznyakov are supported by the EU project ACSI (FP7-ICT-257593).

References

1. Grau, B.C., Kharlamov, E., Zheleznyakov, D.: Ontology contraction: Beyond propositional paradise. In: Proc. Alberto Mendelzon International Workshop on Foundations of Data Management (AMW). (2012)
2. Grau, B.C., Jimenez-Ruiz, E., Kharlamov, E., Zheleznyakov, D.: Ontology evolution under semantic constraints. In: Proc. of International Conference on Principles of Knowledge Representation and Reasoning (KR), Rome, Italy (2012)
3. Horrocks, I., Patel-Schneider, P.F., van Harmelen, F.: From $\mathcal{SHIQ}$ and RDF to OWL: the making of a web ontology language. In Journal of Web Semantics **1**(1) (2003) 7–26
4. Cuenca Grau, B., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., Sattler, U.: OWL 2: The next step for OWL. In Journal of Web Semantics **6**(4) (2008) 309–322
5. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F., eds.: The Description Logic Handbook: Theory, Implementation, and Applications. In Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F., eds.: Description Logic Handbook, Cambridge Uni. Press (2003)

6. Flouris, G., Manakanatas, D., Kondylakis, H., Plexousakis, D., Antoniou, G.: Ontology change: classification and survey. Knowledge Engineering Review **23**(2) (2008) 117–152
7. Fridman Noy, N., Kunnatur, S., Klein, M., Musen, M.: Tracking changes during ontology evolution. In: Proc. of International Semantic Web Conference (ISWC). Volume 3298 of LNCS., Springer (2004) 259–273
8. Haase, P., van Harmelen, F., Huang, Z., Stuckenschmidt, H., Sure, Y.: A framework for handling inconsistency in changing ontologies. In: International Semantic Web Conference (ISWC). (2005) 353–367
9. Haase, P., Stojanovic, L.: Consistent evolution of OWL ontologies. In: Proc. of Extended Semantic Web Conference (ESWC). (2005)
10. Qi, G., Du, J.: Model-based revision operators for terminologies in Description Logics. In: Proc. of International Joint Conference on Artificial Intelligence. (2009) 891–897
11. Calvanese, D., Kharlamov, E., Nutt, W., Zheleznyakov, D.: Evolution of DL-Lite Knowledge Bases. In: Proc. of International Semantic Web Conference (ISWC), Shanghai, China (2010)
12. Wang, Z., Wang, K., Topor, R.W.: Revising general knowledge bases in Description Logics. In: Proc. of International Conference on Principles of Knowledge Representation and Reasoning (KR). (2010)
13. Jimenez-Ruiz, E., Cuenca Grau, B., Horrocks, I., Berlanga, R.: Supporting concurrent ontology development: Framework, algorithms and tool. Data and Knowledge Engineering **70:1** (2011)
14. Konev, B., Walther, D., Wolter, F.: The logical difference problem for description logic terminologies. In: Proc. of International Joint Conference on Automated Reasoning (IJCAR). (2008)
15. Gonçalves, R.S., Parsia, B., Sattler, U.: Analysing the evolution of the NCI thesaurus. In: Proc. of CBMS. (2011) 1–6
16. Alchourrón, C.E., Gärdenfors, P., Makinson, D.: On the logic of theory change: Partial meet contraction and revision functions. In Journal of Symbolic Logics **50**(2) (1985) 510–530
17. Peppas, P.: Belief revision. In F. van Harmelen, V.L., Porter, B., eds.: Handbook of Knowledge Representation. (2007)
18. Hartung, M., Kirsten, T., Rahm, E.: Analyzing the evolution of life science ontologies and mappings. In: Proc. of Data Integration in The Life Sciences workshop (DILS). Volume 5109 of LNCS., Springer (2008) 11–27
19. Kalyanpur, A., Parsia, B., Sirin, E., Grau, B.C.: Repairing unsatisfiable concepts in OWL ontologies. In: Proc. of Extended Semantic Web Conference (ESWC). (2006) 170–184
20. Kalyanpur, A., Parsia, B., Sirin, E., Hendler, J.A.: Debugging unsatisfiable classes in OWL ontologies. In Journal of Web Semantics **3**(4) (2005) 268–293
21. Schlobach, S., Huang, Z., Cornet, R., van Harmelen, F.: Debugging incoherent terminologies. Journal of Automated Reasoning **39**(3) (2007) 317–349
22. Kharlamov, E., Zheleznyakov, D.: Capturing instance level ontology evolution for DL-Lite. In: Proc. of International Semantic Web Conference (ISWC), Bonn, Germany (2011)
23. Wang, Z., Wang, K., Topor, R.W.: A new approach to knowledge base revision in DL-Lite. In: Proc. of Conference on Artificial Intelligence (AAAI). (2010)
24. Lenzerini, M., Savo, D.F.: On the evolution of the instance level of DL-Lite knowledge bases. In: Description Logics. (2011)