Comparing ontologies with *ecco*

Rafael S. Gonçalves, Bijan Parsia, and Uli Sattler

School of Computer Science, University of Manchester, Manchester, United Kingdom

Abstract. In this paper we present the diff tool *ecco*, which detects changes to both axioms and concepts between OWL ontologies. Furthermore, the tool aligns axiom changes between each other, according to a fine-grained change categorisation, and subsequently aligns axiom changes with the concepts that each of those directly affect. The diff is open source, and made available as a standalone command-line tool, as well as a Web-based application.

1 Introduction

The diff tool presented in this paper, *ecco*, incorporates structural and semantic techniques to detect differences between OWL ontologies at both axiom and concept level.

At the axiom level, *ecco* uses structural difference to detect additions and removals, and subsequently verifies whether these changes have any logical impact (i.e., whether they are *logically effectual* or *ineffectual*). Based on these two, coarse-grained categories, we derive finer-grained ones that reflect the apparent impact of the changes detected, as described in [1]. For instance, by further constraining an axiom $A \sqsubseteq B$ into $A \sqsubseteq B \sqcap C$ we "strengthen" it, and the relation between the stronger axiom and its preceding version is made explicit by our categorisation (that is, we align the source and target of the change), and suitably presented by our tool. Such a categorisation of changes is shown to facilitate the navigation through, and analysis of axioms in the diff.

In addition to detecting axiom changes, *ecco* computes entailment and term differences between ontologies. Differences at the entailment level are computed according to several entailment grammars explained in [2], and only shown to users upon request. These entailment differences are used to retrieve the sets of concepts that were specialised or generalised, that is, concepts which have a new superconcept or a new subconcept, respectively. Finally, the tool aligns concept changes with axiom changes, where each (effectual) axiom is aligned with the concepts it directly affects.

The diff *ecco* is freely available as a command-line tool with advanced features, as well as a Web-based application. Both of these output an XML change set file and a transformation of that into HTML, which allows users to browse through and focus on those changes of utmost interest using any Web browser and operating system.

2 Related Work

Structural difference, based on OWL's notion of structural equivalence, is used in several tools to present axiom changes between ontologies; specifically within ContentCVS [3], Bubastis [6] and OWLDiff [5]. However, none of these produce any form of alignment between axioms. The tool ContentCVS also computes entailment differences between ontologies, but does not extrapolate affected concepts from the entailments in the diff. The tool CEX [4] computes entailment and concept differences between acyclic $\mathcal{EL}$ terminologies with role hierarchies and range restrictions. In addition to the major restriction on its input, CEX does not compute changes between axioms, nor does it produce any form of alignment between axioms and the terms they affect.

3 *ecco*: A hybrid diff for OWL 2 ontologies

The diff *ecco* is an open source Java tool available at `https://github.com/rsgoncalves/ecco`. Most major (i.e., computationally expensive) operations are performed in parallel, taking advantage of new concurrency features in Java 7. The command line interface allows tuning the diff using advanced options, all of which can naturally be used programatically as well. Additionally, there is a Web-based front end that allows users to use the system on small to medium ontologies, without having to download it. A demo instance of the Web-based version of *ecco* is deployed at `http://owl.cs.manchester.ac.uk/diff`.[1] In order to demonstrate the functionality of the tool, as well as how its output can be interpreted, we carry out an example diff walkthrough using the toy ontologies in Table 1, and further on we show the output of *ecco* on those same ontologies.

Table 1: Example ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$.

	$\mathcal{O}_1$		$\mathcal{O}_2$
α_1 :	$A \sqsubseteq B$	β_1 :	$A \sqsubseteq B$
α_2 :	$B \sqsubseteq C$	β_2 :	$B \sqsubseteq C \sqcap F$
α_3 :	$C \sqsubseteq \exists r.X$	β_3 :	$C \sqsubseteq \exists r.X$
α_4 :	$\exists r.X \sqsubseteq \exists r.Y$	β_4 :	$X \sqsubseteq D$
α_5 :	$X \sqsubseteq D \sqcap E$	β_5 :	$F \sqsubseteq \exists r.Y \sqcap G$
α_6 :	$F \sqsubseteq \exists r.Y$		

To start with, *ecco* computes the sets of additions and removals between $\mathcal{O}_1$ and $\mathcal{O}_2$ according to structural equivalence. From these changes the tool distinguishes between those that have logical impact (effectual) and those that do not (ineffectual), that is, we check which removed axioms are entailed by $\mathcal{O}_2$ (ineffectual removals), and analogously for added axioms. This coarse-grained categorisation is shown in Table 2.

Table 2: Coarse-grained categorisation of changes in $\text{diff}(\mathcal{O}_1, \mathcal{O}_2)$.

Removals		**Additions**	
Effectual	**Ineffectual**	**Effectual**	**Ineffectual**
$\{\alpha_4, \alpha_5\}$	$\{\alpha_2, \alpha_6\}$	$\{\beta_2, \beta_5\}$	$\{\beta_4\}$

[1] The code is hosted at `https://github.com/rsgoncalves/ecco-webui`.

Subsequently, *ecco* performs a fine-grained categorisation of the changes according to entailment and justification relations, allowing us to identify and align changes between ontologies. The categorisation is shown in Table 3, where we denote effectual additions as $\mathrm{EffAdds}(\mathcal{O}_1, \mathcal{O}_2)$, ineffectual additions as $\mathrm{IneffAdds}(\mathcal{O}_1, \mathcal{O}_2)$, and analogously for removals.

Table 3: Fine-grained categorisation of changes in $\mathrm{diff}(\mathcal{O}_1, \mathcal{O}_2)$.

	Coarse-grained category	**Fine-grained category**	**Axiom change**	**Axiom alignment**
Additions	$\mathrm{EffAdds}(\mathcal{O}_1, \mathcal{O}_2)$	$\mathrm{Strgth}(\mathcal{O}_1, \mathcal{O}_2)$	β_2	$\{\alpha_2\}$
		$\mathrm{StrgthNT}(\mathcal{O}_1, \mathcal{O}_2)$	β_5	$\{\alpha_6\}$
	$\mathrm{IneffAdds}(\mathcal{O}_1, \mathcal{O}_2)$	$\mathrm{NewRed}(\mathcal{O}_1, \mathcal{O}_2)$	β_4	$\{\alpha_5\}$
Removals	$\mathrm{EffRems}(\mathcal{O}_1, \mathcal{O}_2)$	$\mathrm{Weakng}(\mathcal{O}_1, \mathcal{O}_2)$	α_5	$\{\beta_4\}$
		$\mathrm{PrRem}(\mathcal{O}_1, \mathcal{O}_2)$	α_4	–
	$\mathrm{IneffRems}(\mathcal{O}_1, \mathcal{O}_2)$	$\mathrm{NewRed}(\mathcal{O}_1, \mathcal{O}_2)$	α_2 α_6	$\{\beta_2\}$ $\{\beta_5\}$

The fine-grained categorisation shown in Table 3 reveals such changes as α_2 being strengthened into β_2 with shared terms, and similarly α_6 into β_5 though using new terms. In the set of ineffectual changes we have only new retrospective (resp. prospective) redundancies, that is, added axioms (resp. removed axioms) for which there are more constraining axioms in $\mathcal{O}_1$ (resp. $\mathcal{O}_2$). For instance, β_4 is a weaker version of α_5. Within the effectual removals we have α_5 which is weakened into β_4 using shared terms, and α_4 which has no identifiable relation with axioms in $\mathcal{O}_2$.

After axiom changes are categorised, *ecco* detects which atomic terms had their meaning affected between ontologies.[2] In order to do that, *ecco* first computes differences w.r.t. finite sets of entailments, defined according to some entailment grammar.[3] From the sets of lost and gained entailments, so called *witness axioms*, the tool extrapolates the affected terms depending on whether these occur on the left hand side of an entailment difference (specialised term), or right hand side (generalised term). Furthermore, *ecco* distinguishes between whether a concept A is directly affected, or indirectly affected via some other concept change which propagates to A. Finally, the tool aligns term and axiom changes based on witness axioms and their justifications: if the justification for a witness axiom that witnesses a direct (resp. indirect) change to A contains an effectual change α, then α is said to directly (resp. indirectly) affect A, denoted $\{\alpha\} \rightarrow^d A$ (resp. $\{\alpha\} \rightarrow^i A$). In Table 4 we show the result of computing and aligning term differences according to entailments between atomic concepts.

---

[2] Currently restricted to concept changes only, though roles are easily added.

[3] The tool supports all entailment grammars specified in [2], e.g., differences w.r.t. atomic subsumptions, or differences over subsumptions involving asserted subconcepts.

Table 4: Affected concepts with corresponding witness axioms and justifications.

	Affected concept	**Effect**	**Witness axioms**	**Justification(s)**	**Axiom alignment**
Specialised	A	Gained superconcept	$A \sqsubseteq F$ $A \sqsubseteq G$	$\{\beta_1, \beta_2\}$ $\{\beta_1, \beta_2, \beta_5\}$	$\{\beta_2\} \rightarrow^i A$ $\{\beta_5\} \rightarrow^i A$
	B	Gained superconcept	$B \sqsubseteq F$ $B \sqsubseteq G$	$\{\beta_2\}$ $\{\beta_2, \beta_5\}$	$\{\beta_2\} \rightarrow^d B$ $\{\beta_5\} \rightarrow^i B$
	X	Lost superconcept	$X \sqsubseteq E$	$\{\alpha_5\}$	$\{\alpha_5\} \rightarrow^d X$
Generalised	G	Gained subconcept	$F \sqsubseteq G$ $B \sqsubseteq G$ $A \sqsubseteq G$	$\{\beta_5\}$ $\{\beta_2, \beta_5\}$ $\{\beta_1, \beta_2, \beta_5\}$	$\{\beta_5\} \rightarrow^d G$ $\{\beta_2\} \rightarrow^i G$ — ” —
	F	Gained subconcept	$B \sqsubseteq F$ $A \sqsubseteq F$	$\{\beta_2\}$ $\{\beta_1, \beta_2\}$	$\{\beta_2\} \rightarrow^d F$ — ” —
	E	Lost subconcept	$X \sqsubseteq E$	$\{\alpha_5\}$	$\{\alpha_5\} \rightarrow^d E$

4 Discussion

In this paper we presented the diff tool *ecco*, and exemplified how its categorisation mechanisms and presentation of differences can facilitate change analysis. By categorising axioms the tool presents the changed axioms and what they are a change of. Thus we can group and align changes according to their impact, allowing users to shift their attention to specific types of changes rather than going through an unstructured change set while inspecting both ontologies. Based on the combination and alignment of axiom and term changes, *ecco* provides a comprehensive and intelligible change report that would suit most ontology engineers, whether they are only interested in term or axiom changes. In particular, it facilitates the understanding of the impact of axiom changes on the ontology, and the meaning of its terms (via entailment differences). The diff tool is freely distributed as both a command line tool and a Web-based application.

References

1. Gonçalves, R.S., Parsia, B., Sattler, U.: Categorising logical differences between OWL ontologies. In: Proc. of CIKM-11 (2011)
2. Gonçalves, R.S., Parsia, B., Sattler, U.: Concept-based semantic difference in expressive description logics. In: Proc. of ISWC-12 (2012)
3. Jiménez-Ruiz, E., Cuenca Grau, B., Horrocks, I., Berlanga Llavori, R.: Supporting concurrent ontology development: Framework, algorithms and tool. DKE 70(1), 146–164 (2011)
4. Konev, B., Ludwig, M., Wolter, F.: Logical difference computation with CEX 2.5. In: Proc. of IJCAR-12 (2012)
5. Křemen, P., Šmíd, M., Kouba, Z.: OWLDiff: A practical tool for comparison and merge of OWL ontologies. In: Proc. of DEXA-11 (2011)
6. Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., Zhukova, A., Brazma, A., Parkinson, H.: Modeling sample variables with an experimental factor ontology. Bioinformatics 26(8), 1112–1118 (2010)