Focus Set Semantic Differences

Ghadah Alghamdi*
Department of Computer Science,
Dar Al-Hekma University
Kingdom of Saudi Arabia
gghamdi@dah.edu.sa

Renate A. Schmidt
Department of Computer Science,
The University of Manchester
United Kingdom
renate.schmidt@manchester.ac.uk

Yongsheng Gao
SNOMED International
United Kingdom
yga@snomed.org

ABSTRACT

Ontologies have been utilized widely as sources for formally organized information in a range of fields. SNOMED CT ontology is used in the (bio-)medical field because it offers a comprehensive multilingual vocabulary for encoding different domains of electronic health records and clinical knowledge, resulting in a very large and complex ontology. SNOMED CT is regularly updated to reflect domain changes that are often significant. Ontology engineers are often interested in understanding changes to the meaning (semantics) of an ontology for a variety of reasons, such as providing a mechanism for an accurate and safe alignment, integration, or reuse of an ontology version. Computing the semantic difference between SNOMED CT versions tends to produce rather large changes that are difficult to analyze. To address this problem, this paper presents a method for tracking semantic differences between large-scale ontologies that produce concentrated sets of semantic differences relating to input focus symbols chosen by the user. Our solution circumvents the size issue by initially constructing subontologies for the focus symbols in question. The resulting differences are related to the meaning of focus concept definitions for specific ontology subdomain, where some of these differences would not have been generated without this focused method for identifying semantic differences between ontologies. A case study using SNOMED CT has shown the proposed approach is useful for domain experts.

CCS CONCEPTS

• **Computing methodologies → Description logics**; **Ontology engineering**; • **Theory of computation → Abstraction**.

KEYWORDS

Semantic differences, Subontologies, Forgetting/ Uniform Interpolation, Ontology Engineering, SNOMED CT

ACM Reference Format:
Ghadah Alghamdi, Renate A. Schmidt, and Yongsheng Gao. 2023. Focus Set Semantic Differences. In *Proceedings of the 12th Knowledge Capture Conference (K-CAP '23), December 5–7, 2023, Pensacola, Florida, USA*. ACM, New York, NY, USA, 8 pages. https://doi.org/X

*Ghadah Alghamdi is the corresponding author of this paper.

K-CAP '23, December 5–7, 2023, Pensacola, Florida, USA

ACM ISBN X…$15.00
https://doi.org/X

1 INTRODUCTION

An ontology is a formal description of knowledge given as a set of definitions/axioms about a certain domain expressed using the Web Ontology Language (OWL) [36]. The underlying formalism of OWL ontologies is Description Logic (DL), which formalizes the relationships between the ontology's concepts [4]. Ontologies are used in a wide range of fields, such as (bio-)medicine [33], engineering [30], and law [10].

In the (bio-)medical fields, the National Cancer Institute Thesaurus (NCIt) ontology [7], the Systemized Nomenclature of Medicine - Clinical Terms (SNOMED CT) ontology [32], and the Gene Ontology (GO) [34] are prominent. These ontologies have been developed utilizing principled design and deployment techniques, typically by teams of engineers and over many years.

SNOMED CT is the most comprehensive, multi-lingual bio-medical ontology in the world for standardizing clinical terminology globally. It is built on a multi-hierarchical taxonomy with approximately 350K concepts that describe several subdomains of biomedicine, such as clinical symptoms, diseases, operations, body structures, and pharmaceuticals. The most recent version of SNOMED CT (May 2023) contains 126 roles, 361 044 concepts, and 362 735 axioms.

SNOMED International[1] builds and maintains the core SNOMED CT ontology (International edition), which is used as the basis for building e.g., country extensions. Additionally, it maintains a repository of reference sets (refsets) [13]. These reference sets allow for the creation of SNOMED CT content subsets relevant to specific clinical areas of information.

SNOMED CT is regularly updated by adding new concepts, retiring outdated concepts, and modifying existing concepts. In particular, the International edition is released every month to reflect domain changes that are often significant. For instance, the content of SNOMED CT had almost 192 000 changes in the releases of January 2015, July 2015, and January 2016 [28].

SNOMED International offers a way to examine potential changes in each ontology release. Every release comes with a series of delta files that detail how the current version differs from the previous one [14]. Such diff reports are produced based on structural differences. Because these reports only detail structural changes, relying solely on them may leave out significant consequences of updating the ontology. Typically, ontology engineers are interested in understanding the changes to the meaning (semantics) of the ontology regardless of its syntactic structure. The diff reports do not reflect whether one version of the ontology is semantically different from another. Tracking semantic differences is critical for various reasons: (i) checking that the new version's changes are safe in the sense that the new version is a conservative extension of the previous one [6, 24], (ii) identifying unexpected entailments

[1] https://www.snomed.org/

as a result of changing the ontology, and (iii) providing means for the alignment, integration, or reusing of an ontology version in a correct and safe way [15, 31].

Computing semantic differences Diff($\mathcal{O}_1$,$\mathcal{O}_2$) between two ontology versions $\mathcal{O}_1$ and $\mathcal{O}_2$ is done by computing all of the logical entailments $\mathcal{V}_2$ of $\mathcal{O}_2$, and then collect from $\mathcal{V}_2$ the axioms that are not entailed by $\mathcal{O}_1$. Computing $\mathcal{V}_2$ may however not be feasible as $\mathcal{V}_2$ is in general infinite [18]. Rather than computing all of the logical entailments of $\mathcal{O}_2$, [23] defined semantic difference between two ontologies as computing the strongest entailments that are not entailed by the other ontology. These strongest entailments can be computed using an extraction technique known as uniform interpolation (UI). UI generates views of the source ontology that semantically preserves all of the views' logical entailments over a restricted signature.

Other methods to compute semantic difference were introduced in [9] and [17]. ECCO [9] uses a hybrid approach to identify differences in both structure and meaning. The tool's ability to identify semantic differences mainly relies on a standard reasoner. On the other hand, CEX [17] detects semantic difference using the notion of insperability. The tool returns a list of concept names A that are involved in a list of semantic differences of the form $A \sqsubseteq C$ or $C \sqsubseteq A$, but it does not return the whole axiom including the affected concept. Both tools may fail to reveal some of the strongest inferred entailments. However, the method introduced in [22] is capable of computing all of the strongest entailments as it uses uniform interpolation to track semantic differences between ontologies.

Existing methods for tracking semantic differences are beneficial and can in principle be used to compare ontologies as large as SNOMED CT. It has however been found that the resulting witness sets are rather big and difficult to analyze. For example, the case study in [37] demonstrated that the July 2017 International edition of SNOMED CT did not entail over 8 400 axioms from the January 2017 version, whereas the Australian 2017 extension did not entail nearly 43 000 axioms from the July 2017 International edition. Such large sets are difficult to analyze.

The size issue is also present in the *delta release* files used in the SNOMED community, where the amount of generated differences are not only large but also dispersed, i.e., belonging to several ontology subhierarchies [27]. Users are overwhelmed by the enormous number of changes to analyze without the ability to restrict such differences to certain subdomains of the ontology.

The solution suggested in this paper is to use ontology extraction to restrict semantic difference computation to corresponding extracts of the two ontologies. Syntactic Locality Based Modularisation (SLBM) is an ontology extraction method that returns an extract consisting of a subset of the axioms of an ontology for a given seed signature Σ. Evaluations in [3, 5, 35] have revealed a very large number of symbols of the computed modules that are outside the input signature Σ. On the other hand, the UI method gives an extract where their symbols are exactly in the range of the input signature [19, 23, 26]. However, because of how the method works, some axioms can lose their definitions because symbols of such definitions were not specified as input signatures to the method. This can result in very small extracts that do not contain full definitions of symbols in Σ [3]. Additionally, UIs are expensive to compute and termination is not always guaranteed [25].

In [3], we introduced a method to compute subontologies from large $\mathcal{ELH}$ ontologies. The subontology generation method meets the requirements of SNOMED CT users by providing complete semantics for the definition of input focus symbols while remaining concise and adhering to SNOMED CT modeling standards.

Currently, there are no tools for tracking semantic differences between *subdomains* of SNOMED CT. In this paper, we combine two recently introduced techniques to compute witnesses (aka semantic differences) for subdomains of the ontology: the subontology generation method [1], and the UI-Diff method [22]. The idea of our method, presented in this paper, is to produce a subontology for a specified focus set of two different $\mathcal{ELH}$ terminologies' versions based on the user's selection, and then utilize the generated subontologies to compute witnesses between them. The generated witnesses are then categorized using our witness analysis scheme in order to facilitate witness examination.

The paper is organized as follows. Section 2 gives preliminary definitions and background information. Section 3 describes the subontology generation and UI-Diff methods utilized to build our method. Section 4 details our method of generating the semantic differences between subontologies and describes a scheme to analyze them. In Section 5, we present our evaluation results and two cases of using the results for analysis and conclude in Section 6.

2 PRELIMINARIES AND BACKGROUND

Let $\mathsf{N_C}$ and $\mathsf{N_R}$ be disjoint sets of concept and role names respectively. The union of such sets forms the signature of an ontology $\mathcal{O}$. The signature $\mathrm{sig}(\xi)$ of a syntactic object or ontology ξ is a set of concept and role names that occur in ξ. By $\mathrm{sig}_C(\xi)$ we denote the set of concept names that occur in ξ. The set of *$\mathcal{EL}$-concepts* C and D, and the sets of *$\mathcal{ELH}$-axioms* α are built according to the grammar rules: $C, D ::= A \mid C \sqcap D \mid \exists r.C$ and $\alpha ::= C \sqsubseteq D \mid C \equiv D \mid r \sqsubseteq s$, where $A \in \mathsf{N_C}$ and $r, s \in \mathsf{N_R}$.[2] An *$\mathcal{ELH}$-TBox* is a finite set of $\mathcal{ELH}$-axioms.

Concepts are interpreted as sets, roles are binary relations, $\sqsubseteq$ as a subset relation, $\equiv$ as a set equality and $\sqcap$ as an intersection. Intuitively $\exists r.C$ is the set of elements that are r-related to some element in C [4].

A *terminology* is a TBox that contains only axioms of the form $A \equiv C$ or $A \sqsubseteq C$, with A appearing not more than once on the left-hand side of an axiom. If in a terminology A does not depend on itself, the terminology is *acyclic*. An $\mathcal{ELH}$-terminology is *normalized* iff it only contains axioms of the forms $A \sqsubseteq C$, $A \sqsubseteq \exists r.C$, and $\exists r.C \sqsubseteq A$, where $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and C is an $\mathcal{EL}$-concept.

In SNOMED CT, existential restrictions $\exists r.C$ existing in a concept definition A are concepts that define a certain characteristic of that concept, hence they are also called *defining characteristics*. Listing all of the defining characteristics of a concept facilitates implementation, recording, storage, and retrieval within SNOMED CT [33]. SNOMED CT's earlier releases (2017 and earlier) were stated as $\mathcal{ELH}$ terminologies. As of the January 2018 release, SNOMED CT began to support higher expressive language constructors, including role chains, transitive, and reflexive roles, as well as GCI axioms; $C \sqsubseteq A$ where C is an $\mathcal{EL}$-concept and A is a concept name.

[2] These grammar rules are sufficient for the $\mathcal{ELH}$ fragment we are considering.

A *justification* $\mathcal{J}$ of an implied axiom α in an ontology O is a $\subseteq$-minimal subset of axioms of O such that $\mathcal{J}$ implies α, that is $\mathcal{J} \subseteq O$, $\mathcal{J} \models \alpha$ and there is no $\mathcal{J}' \subset \mathcal{J}$ such that $\mathcal{J}' \models \alpha$ [16].

3 UTILIZED METHODS

This section describes the methods utilized in our approach.

3.1 The UI-Diff Method

One of the most significant considerations for an ontology engineer while modifying an existing ontology is ensuring that the provided modifications do not interfere with the meaning of the concepts/roles beyond the fragment of the ontology being edited [23]. Semantic diffing methods are useful in such situations.

The UI-Diff method uses uniform interpolation, which aids in forgetting the symbols of axioms that are not relevant for examination. Forgetting these symbols computes inferred axioms, which are then checked to see whether or not the other ontology version entails them. This way it can be guaranteed that the introduced changes are safe for the ontology and do not contradict what is already stated in the ontology [23]. It is emphasized in [18] that the interesting differences are those that are expressed in their shared signature, not in the symbols used in only one of the two ontology versions. The use of forgetting is appropriate for such applications as it helps eliminate those symbols that are included in only one of the two ontology versions.

In [22], the UI-Diff method was used to track semantic differences between ontologies as large as SNOMED CT. Their case study showed that there are over 1M axioms in the 2017 January release and were not entailed by the 2016 July release. Furthermore, while comparing different versions of the NCIt ontology using semantic diff tools like ECCO [9] and CEX [17], a significant number (57K axioms) of differences were discovered [8].

The UI-Diff method can limit the generation of semantic differences to a fewer number by specifying which symbols to include in the common signature set while generating the witnesses. Picking the right common symbols to construct the set of witnesses is however a difficult task [1]. If the user is interested in learning about changes to the definitions of particular concepts of interest and the linkages between their signatures, it is unclear which symbols to pick or exclude that cause such changes.

It is guaranteed that a finite UI-based semantic difference will be detected when using the UI-Diff method to generate witnesses. The difference between O_1 and O_2 is empty iff $O_1 \models \mathcal{UI}_2$, where $\mathcal{UI}_2$ is a Σ-uniform interpolant of O_2 computed for $\Sigma \subseteq \text{sig}(O_1) \cap \text{sig}(O_2)$. If $O_1 \not\models \mathcal{UI}_2$ it means that every $\alpha \in \mathcal{UI}_2$ not entailed by O_1 is a witness [22].

Definition 3.1 (UI-based Semantic Difference / UI-witness). Let O_1 and O_2 be $\mathcal{ELH}$ ontologies. Let Σ be a subset of the common signature of O_1 and O_2. The UI-based semantic difference between O_1 and O_2 is the set UI-Diff(O_1,O_2) of all $\mathcal{ELH}$-axiom α such that (i) sig(α) $\cap\ \Sigma$, (ii) $\alpha \in \mathcal{UI}_2$ and (iii) $O_1 \not\models \alpha$, where $\mathcal{UI}_2$ is a Σ-uniform interpolant of O_2. An axiom α satisfying these conditions is a *UI-witness* of a difference in O_2 w.r.t. O_1. We denote by $\mathcal{W}$ the set of all such axioms α.

In the use case of ontology versioning, computing UI-Diff(O_1, O_2) results in witnesses indicating new information from O_1 added to O_2. These witnesses, which are referred to as $\mathcal{W}_2$, are axioms that are entailed by O_2 but not O_1. When O_1 and O_2 are switched in UI-Diff, lost information $\mathcal{W}_1$ is computed with the assumption that O_2 is the more recent version while O_1 is the older version.

To compute UI-Diff(O_1, O_2), two main steps should be performed:

- Compute $\mathcal{UI}_2$ of O_2 for Σ = sig(O_1) $\cap$ sig(O_2) using the UI method.
- Compute the set $\mathcal{W}_2$, which consists of the axioms $\alpha \in \mathcal{UI}_2$ but $O_1 \not\models \alpha$ using an external DL reasoner.

3.2 The Subontology Extraction Method

In [3], we introduced a method that extracts *subontologies* from the source ontology expressed in $\mathcal{ELH}$ for input focus concepts describing part of the source ontology's domain. The method is based on the principle of abstracted definitions, which helps include only what is truly necessary in the resulting subontology. The idea of abstracted definitions is to compute definitions for focus concepts that are based on the focus concept's closest primitive ancestor(s). The abstracted definitions should also include inferred existential restrictions that are required to complete the definition upon abstraction. Abstracting the definitions based on the closest primitive concept helps ensure that all the defining characteristics of the focus concept are included in its definition (see Property 2 in [1, 3]).

The method generates a subontology $\mathcal{S}$ for a focus set Σ_F extracted from an ontology O. It ensures the inclusion of full definitions of concepts in the focus set Σ_F in the form of abstracted definitions. The signature of focus concepts' definitions is referred to as the *supporting set* Σ_S defined as Σ_S= sig($\mathcal{S}$) $\setminus \Sigma_F$. If subsumption relationships exist between supporting set symbols Σ_S in O, then they are added to the subontology $\mathcal{S}$.

Definition 3.2 (Focus Set Subontology). Let O be an $\mathcal{ELH}$ ontology and Σ_F a focus set of concept and role names. $\mathcal{S}$ is a *focus set subontology* of O for Σ_F if the following conditions are satisfied: (i) $\Sigma_F \subseteq$ sig($\mathcal{S}$); (ii) for every $\mathcal{ELH}$-axiom α where sig(α) $\subseteq$ sig($\mathcal{S}$) we have: (a) If $\mathcal{S} \models \alpha$ then $O \models \alpha$, and (b) if α is of the form $A \sqsubseteq B$ or $r \sqsubseteq s$, then $\mathcal{S} \models \alpha$ when $O \models \alpha$, where A and B are concept names, and r and s are role names.

Notably, if O is an acyclic $\mathcal{ELH}$-terminology, then the axiom α of a focus concept name in Σ_F have the same logical entailments as their definitions in O, i.e., are equivalent to their axioms in O [1].

Axioms within $\mathcal{S}$ can be segmented into two main sets, *focus set axioms* and *supporting set axioms. Focus set axioms* contain possibly abstracted definitions of the form $A \equiv C$ or $A \sqsubseteq C$ where A is a focus concept, while *supporting set axioms* are of the form $B \sqsubseteq C$, where B is a supporting concept. The method only preserves the necessary conditions of the supporting concept as long as the signature of C is in the signature of the focus set axioms.

4 COMPUTING FOCUS SET SEMANTIC DIFFERENCES

The benefit of our approach is demonstrated in the following example.

Example 4.1. Let the ontology O_1 consists of the following axioms:

Renal artery stenosis $\equiv$ Disease $\sqcap$ $\exists$location.Renal artery,
Acute renal failure syndrome $\equiv$ Disease $\sqcap$ $\exists$location.Kidney,
Renal artery $\sqsubseteq$ Vascular structure of kidney,
Vascular structure of kidney $\sqsubseteq$ Kidney

and assume O_2 consists of the axioms:

Renal artery stenosis $\equiv$ Disease $\sqcap$ $\exists$location.Renal artery,
Acute renal failure syndrome $\equiv$ Disease $\sqcap$ $\exists$location.Kidney,
Renal artery $\sqsubseteq$ Arterial supply,
Arterial supply $\sqsubseteq$ Vascular structure of kidney,
Vascular structure of kidney $\sqsubseteq$ Abdominal organ

Assume that the user is interested in tracking the differences between O_1 and O_2 w.r.t. the focus set Σ_F = {**Renal artery stenosis**, **Acute renal failure syndrome**} displayed in bold face. Then one can generate the subontologies $\mathcal{S}_1$ and $\mathcal{S}_2$ for Σ_F from O_1 and O_2 respectively, which are:

Renal artery stenosis $\equiv$ Disease $\sqcap$ $\exists$location.Renal artery,
Acute renal failure syndrome $\equiv$ Disease $\sqcap$ $\exists$location.Kidney,
Renal artery $\sqsubseteq$ Kidney

and

Renal artery stenosis $\equiv$ Disease $\sqcap$ $\exists$location.Renal artery,
Acute renal failure syndrome $\equiv$ Disease $\sqcap$ $\exists$location.Kidney

respectively. $\mathcal{S}_1$ and $\mathcal{S}_2$ are concise subontologies of O_1 and O_2 which contain the information relevant to the focus concepts Σ_F. Computing the UI-Diff($\mathcal{S}_2, \mathcal{S}_1$) to get the axioms entailed by $\mathcal{S}_1$ but not $\mathcal{S}_2$ gives the witness:

$$\{\text{Renal artery} \sqsubseteq \text{Kidney}\}$$

On the other hand, computing the UI-Diff(O_2, O_1) gives the witness:

$$\{\text{Vascular structure of kidney} \sqsubseteq \text{Kidney}\},$$

which is correct but does not tell the user that *Renal artery* is subsumed by *Kidney*.

We notice the resulting witness in UI-Diff($\mathcal{S}_2, \mathcal{S}_1$) is more relevant to the focus set Σ_F than UI-Diff(O_2, O_1). This is because the change identified in UI-Diff($\mathcal{S}_2$, $\mathcal{S}_1$) is about the concept *Renal artery*, which is a supporting concept used in the definition of the focus concept *Renal artery stenosis*. The set UI-Diff(O_2, O_1) does not identify this change. One can argue that we can restrict the common signature when computing UI-Diff(O_2, O_1) by eliminating/forgetting the concept *Vascular structure of kidney* from it to give exactly the same result as when computing UI-Diff($\mathcal{S}_2, \mathcal{S}_1$). However, there are several issues with this approach, some of them are:

1. In general, the user cannot be expected to have enough experience to know which concepts to pick for the input set Σ, when computing the UI-Diff(O_2, O_1).
2. The computation of UI-Diff directly between very large ontologies might not terminate in reasonable time.

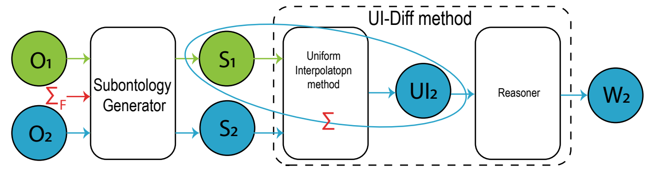

Figure 1: Computing UI-based semantic differences between two ontologies O_1 and O_2 for input focus set Σ_F based on subontology generation

3. Since O_1 and O_2 are typically significantly larger than their corresponding subontologies $\mathcal{S}_1$ and $\mathcal{S}_2$ the witnesses in UI-Diff(O_2, O_1) can also be numerous and hard to analyze, which is what the results in [22] show.

The method proposed in this paper computes subontologies first for a set of focus concepts, and then applies UI-Diff between the generated subontologies. By using our method, the process is directed towards just computing the differences that are important to the focus set symbols.

4.1 UI-Diff Between Subontologies

The aim of our method is to identify semantic differences between two ontology versions for a given set of focus symbols Σ_F that is specific to a particular part of the ontologies. Figure 1 shows our method to detecting semantic differences between two versions of an ontology based on a focus set. Our method takes two ontology versions O_1 and O_2 as input, as well as a focus set Σ_F. The method generates two sets $\mathcal{W}_1$ and $\mathcal{W}_2$, which are the sets of focus set semantic differences. These are computed in two steps:

(1) Using the focus set subontology generation method, we generate two subontologies $\mathcal{S}_1$ and $\mathcal{S}_2$ for the input focus set Σ_F from O_1 and O_2.
(2) Using the UI-Diff method, witnesses between the subontologies $\mathcal{S}_1$ and $\mathcal{S}_2$ are computed. The UI-Diff method in Figure 1 generates axioms α that are in $\mathcal{UI}_2$ computed from $\mathcal{S}_2$ but for which $\mathcal{S}_1 \not\models \alpha$. This phase generates the $\mathcal{W}_2$ witness set. $\mathcal{W}_2$ is a witness set that represents UI-Diff($\mathcal{S}_1, \mathcal{S}_2$). By switching the positions of $\mathcal{S}_1$ and $\mathcal{S}_2$, the method generates the set of witnesses $\mathcal{W}_1$.

Assuming that O_1 is the older version and O_2 is the newer version of an ontology, we can see that UI-Diff($\mathcal{S}_1$, $\mathcal{S}_2$) computes the information gained from $\mathcal{S}_1$ to $\mathcal{S}_2$, or the information lost in $\mathcal{S}_2$ from $\mathcal{S}_1$, for a given focus set Σ_F.

After generating the witness sets, we analyze them using the segmentation technique described in the next section.

4.2 Analysis of the Witness Sets

We differentiate between the resulting witnesses in this section based on whether the witness is a stated or inferred axiom. We also differentiate witnesses based on whether the witness is related to a focus or a supporting concept. We provide an analysis method based on such distinctions to analyze the resulting witnesses by splitting them according to the indicated groupings.

Stated or Inferred Witnesses. Distinguishing stated witness axioms from inferred witnesses allows for an unambiguous understanding of the resulting witnesses, as inferred witnesses highlight hidden axioms or unanticipated consequences discovered while tracking semantic differences [37]. The sets of stated and inferred witnesses are defined as follows.

Definition 4.2 (Stated (inferred) Witness). Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two *normalized* ontologies. Let $\mathcal{W}_1$ be the witness set of all axioms α given by computing UI-Diff($\mathcal{O}_2$,$\mathcal{O}_1$). We say that $\mathcal{IW}_1$ is the set of inferred witnesses where $\mathcal{O}_1 \models \alpha$, but $\alpha \notin \mathcal{O}_1$. If $\alpha \in \mathcal{O}_1$ then α is a stated witness. All stated witnesses are denoted by $\mathcal{SW}_1$.

To generate the set of inferred $\mathcal{IW}_1$ and stated $\mathcal{SW}_1$ witnesses, we normalize the ontology $\mathcal{O}_1$ to bring it to a form similar to the axioms in the set $\mathcal{W}_1$. This process yields the normalized ontology $\mathcal{O}_{normalized}$. Then, we check the axioms α in $\mathcal{W}_1$ to see if α is stated in $\mathcal{O}_{normalized}$ or not. If $\alpha \in \mathcal{O}_{normalized}$, then it is stated and added to $\mathcal{SW}_1$, else, it is inferred and added to $\mathcal{IW}_1$. By normalizing the ontology $\mathcal{O}_1$, we correctly identify whether the axiom α is a stated or inferred axiom. Without this normalization process, witness axioms in $\mathcal{W}_1$ that have a structural form different from that in the original ontology $\mathcal{O}_1$ would be incorrectly identified as inferred witnesses. For instance, if a witness axiom α in $\mathcal{W}_1$ is $A \sqsubseteq B$ while the axiom in $\mathcal{O}_1$ is $A \sqsubseteq B \sqcap C$, then $\alpha \notin \mathcal{O}$ and would be incorrectly added to the set of inferred witnesses $\mathcal{IW}_1$.

Focus or Supporting Concept Witnesses. Distinguishing between focus or supporting witnesses aids the analysis process. A change to a focus concept axiom may take precedence over a change to a supporting concept axiom in the analysis process done by users. Furthermore, in the experience of SNOMED CT terminologists, there are cases where the emphasis is on revealing differences affecting the subsumption relationships between supporting symbols that exist in the definitions of focus concepts, which are not provided by existing tooling [2]. We give an example from the evaluation results in Section 5.2 showing the use of the supporting set witnesses as a result of segmenting the witness set.

Definition 4.3 (Focus (Supporting) Concept Witness). Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two subontologies extracted from two $\mathcal{ELH}$ terminologies $\mathcal{O}_1$ and $\mathcal{O}_2$ for a focus set Σ_F. Suppose A is in Σ_F (Σ_S) where Σ_S is the supporting set in $\mathcal{S}$. A witness of the form $A \sqsubseteq C$ or $C \sqsubseteq A$ in UI-Diff($\mathcal{S}_1, \mathcal{S}_2$) is said to be a focus (supporting) concept witness where C is an $\mathcal{EL}$-concept. This witness is said to be associated with the focus (supporting) concept name A. Focus concept witnesses will be denoted by $\mathcal{W}_{\Sigma_F}$, while supporting concept witnesses will be denoted by $\mathcal{W}_{\Sigma_S}$.

Figure 2 depicts our witness analysis scheme. The process of analysis starts with establishing whether a witness is a focus concept witness or a supporting concept witness (Definition 4.3). This yields two sets: $\mathcal{W}_{\Sigma_F}$, which denotes witnesses for the focus set, and $\mathcal{W}_{\Sigma_S}$, indicating witnesses for the supporting set. Each focus and supporting concept witness is further divided into two categories based on whether the witness is stated or inferred (see Definition 4.2). As a result, four sets are formed: $\mathcal{SW}_{\Sigma_F}$, $\mathcal{SW}_{\Sigma_S}$, $\mathcal{IW}_{\Sigma_F}$, and $\mathcal{IW}_{\Sigma_F}$. Each of these four sets is further divided by the two axiom forms: $A \sqsubseteq C$ and $C \sqsubseteq A$, where A is a Σ_F or Σ_S concept and C is an $\mathcal{EL}$-concept. The total number of sets produced is eight. We also show possible intersections of the generated sets. An intersection occurs as a result of segmenting the witness sets into focus and supporting concept witnesses, with the left and right-hand sides of axioms being either a focus or a supporting concept. The black oval shape in Figure 2 demonstrates where there may be non-empty intersections, resulting in the sets (1.1.1 ∩ 2.1.2), (1.1.2 ∩ 2.1.1), (1.2.1 ∩ 2.2.2), and (1.2.2 ∩ 2.2.1).

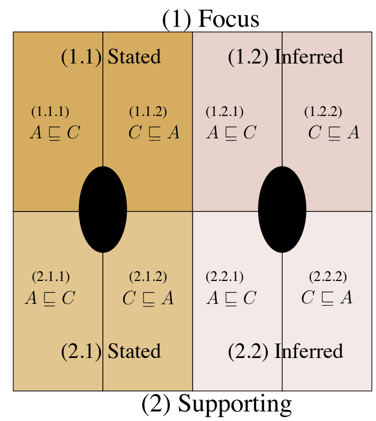

Figure 2: Witness Segmentation Scheme

4.3 Properties of the Generated Witnesses

Using our method, all possible witnesses associated with focus concepts are generated by employing subontologies for tracking semantic differences. This is due to the fact that subontologies preserve the semantic connections associated with the input focus set, which is especially important when the input ontology is a terminology. On the other hand, some witnesses associated with supporting concepts might be discarded by the subontology generation method (cf. Section 3.2). Additionally, as a property of the subontology generation method, all of the defining characteristics are given in the generated abstracted definitions. Such a property is reflected in the generated set of witnesses making the witnesses more relevant to the focus concepts.

PROPERTY 1. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two focus set subnotologies extracted from $\mathcal{O}_1$ and $\mathcal{O}_2$ for a focus set Σ_F respectively, where $\mathcal{O}_1$ and $\mathcal{O}_2$ are acyclic $\mathcal{ELH}$-terminologies. Then UI-Diff($\mathcal{S}_1$,$\mathcal{S}_2$) contains all possible focus concept witnesses $\alpha \in \mathcal{W}_{\Sigma_{F_2}}$ where $\mathcal{S}_1 \not\models \alpha$ including those that represent necessary (and sufficient) conditions of definitions of focus concepts.*

The following example illustrates the property.

Example 4.4. Let $\mathcal{O}_1 = \{A \equiv P_1 \sqcap \exists r.B_1, P_1 \sqsubseteq P_2 \sqcap \exists r.B_2, P_3 \sqsubseteq P_1\}$ and $\mathcal{O}_2 = \{A \equiv P_2 \sqcap \exists r.B_1, P_1 \sqsubseteq P_2 \sqcap \exists r.B_2, P_3 \sqsubseteq P_1\}$ and $\Sigma = \text{sig}(\mathcal{O}_1) = \text{sig}(\mathcal{O}_2)$ are the common symbols of $\mathcal{O}_1$ and $\mathcal{O}_2$. We can see that $\mathcal{O}_1$ and $\mathcal{O}_2$ are similar to each other, with the only difference being in the definition of A. The concept name A in $\mathcal{O}_2$ is defined in terms of P_2 rather than P_1 as in $\mathcal{O}_1$. Computing UI-Diff($\mathcal{O}_2, \mathcal{O}_1$) gives the witness $A \sqsubseteq P_1$. Computing UI-Diff($\mathcal{S}_2, \mathcal{S}_1$) between the generated subontologies $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathcal{O}_1$ and $\mathcal{O}_2$ for the focus set $\Sigma_F = \{A, P_3\}$ respectively, where $= \mathcal{S}_1 = \{A \equiv P_1 \sqcap \exists r.B_1 \sqcap \exists r.B_2, P_3 \sqsubseteq P_1 \sqcap \exists r.B_2, P_1 \sqsubseteq \exists r.B_2\}$ and $\mathcal{S}_2 = \{A \equiv P_2 \sqcap \exists r.B_1, P_3 \sqsubseteq P_1 \sqcap \exists r.B_2, P_1 \sqsubseteq P_2 \sqcap \exists r.B_2\}$ gives the witness set: $\{A \sqsubseteq P_1, A \sqsubseteq \exists r.B_2\}$.

We can see that UI-Diff$(\mathcal{S}_2, \mathcal{S}_1) = \{A \sqsubseteq P_1, A \sqsubseteq \exists r.B_2\}$ includes the additional axiom $A \sqsubseteq \exists r.B_2$. This additional witness was not revealed when comparing between the original ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$. This is because the definition of A in $\mathcal{O}_1$ does not include all the existential restrictions (defining characteristics) that the concept A inherits (including $\exists r.B_2$). Thus, the definition of A in $\mathcal{UI}_1$ generated when computing UI-Diff$(\mathcal{O}_2, \mathcal{O}_1)$ does not include the additional condition ($\exists r.B_2$) that A inherits from P_1. This is because the uniform interpolation method computes only strongest Σ-entailments of the input ontology. For clarification, we display the set of axioms within the $\mathcal{UI}_1$ when computing UI-Diff$(\mathcal{S}_2, \mathcal{S}_1)$. $\mathcal{UI}_1$of UI-Diff$(\mathcal{S}_2, \mathcal{S}_1) = \{A \sqsubseteq P_1, A \sqsubseteq \exists r.B_1, A \sqsubseteq \exists r.B_2, P_1 \sqcap \exists r.B_1 \sqcap \exists r.B_2 \sqsubseteq A, P_3 \sqsubseteq P_1, P_3 \sqsubseteq \exists r.B_2, P_1 \sqsubseteq \exists r.B_1\}$. $\mathcal{UI}_1$ includes normalized axioms of $\mathcal{S}_1$ with their symbols in Σ. After checking which axioms in $\mathcal{UI}_1$ are not entailed by $\mathcal{S}_2$, the resulting set of witnesses is $\{A \sqsubseteq P_1, A \sqsubseteq \exists r.B_2\}$.

5 EVALUATION AND CASE STUDY

This section evaluates our method to computing focus set semantic differences in depth. In our evaluation, we want to: 1) compare the size of the witness sets and 2) analyse the witness sets using our analysis scheme described in Section 4.2.

To conduct the evaluation, we developed a Java prototype implementing our method described in Section 4 using the OWL API [11]. We employed four consecutive versions of the SNOMED CT international edition from January 2016 to July 2017. As focus sets, we used three standard SNOMED CT refsets: the General Practice/Family Practice (GPFP) refset, the International Classification of Nursing Practice Diagnoses (ICNP-Diagnoses) refset, and the International Classification of Nursing Practice Interventions (ICNP-Interventions) refset. The tool and all experimental data used in this study are available at https://tinyurl.com/diff-data.

5.1 Comparing the Size of the Witness Sets

Figure 3 depicts the number of witnesses in comparisons of consecutive SNOMED CT versions. The yellow bar represents the witness sets of tracking the semantic difference between core SNOMED CT editions that were conducted in the evaluation by Liu et al. [22]. The rest of the bars show the number of our findings of tracking the semantic differences between the subontologies computed for the three refsets; GPFP, ICNP, and ICNP-Interventions. Figure 3 does not show the number of witnesses in the comparisons (1607,1601), (1601, 1607) of ICNP-Diagnosis and ICNP-Interventions as they contained zero number of witnesses.

Figure 3 shows that the number of witnesses computed between SNOMED CT's core editions is enormous when compared to the number of witnesses computed for the three refsets. This is to be expected as tracking semantic differences between subontologies computed for the refsets leads to a more focused set of witnesses with a smaller size. For example, the UI-Diff(1601, 1607) comparison of the GPFP refset has just 46 witnesses, compared to 584 witnesses

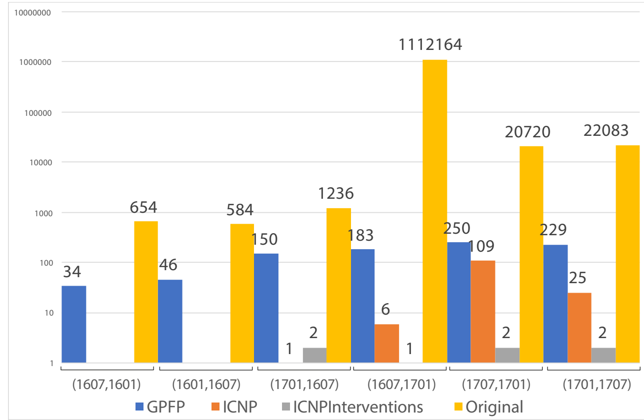

Figure 3: Number of witnesses in various comparisons of subsequent versions of computed subontologies for the SNOMED CT refsets GPFP, ICNP, and ICNP Interventions.

Table 1: Count of total focus and supporting set witnesses

Comparison	(1.1)	(2.1)	(1.1) ∩ (2.1)	(1.2)	(2.2)	(1.2) ∩ (2.2)
GPFP(1607,1601)	21	4	2	11	0	0
GPFP(1601,1607)	35	6	4	4	1	0
GPFP(1701,1607)	124	17	13	20	2	0
GPFP(1607,1701)	157	9	3	17	3	0
GPFP(1707,1701)	148	10	5	90	10	3
GPFP(1701,1707)	157	30	11	51	4	2
ICNP-Diag.(1701,1607)	0	0	0	1	0	0
ICNP-Diag.(1607,1701)	4	0	0	2	0	0
ICNP-Diag.(1707,1701)	102	1	0	0	6	0
ICNP-Diag.(1701,1707)	3	0	0	21	1	0
ICNP-Inter.(1701,1607)	0	0	0	2	0	0
ICNP-Inter.(1607,1701)	1	1	1	0	0	0
ICNP-Inter.(1707,1701)	1	1	0	0	0	0
ICNP-Inter.(1701,1707)	1	0	0	0	1	0

between SNOMED CT core versions. This is also true for the remaining refsets; for example, the witnesses belonging to the ICNP refset in the comparison UI-Diff(1607, 1701) are only a few out of over a million witnesses in the original comparison.

When the number of witnesses across all refsets is compared, the majority of semantic changes are for the SNOMED CT GPFP subontology. This demonstrates that terminologists concentrated their editing operations on the GPFP subsets of the ontology from the January 2016 to July 2017 versions.

5.2 Analysis of the Witness Sets

Table 1 shows statistics on the overall number of focus and supporting set witnesses while tracking semantic difference between different versions of SNOMED CT from the three refsets: GPFP, ICNP, and ICNP-Interventions. These witnesses are divided in accordance with the scheme described in Section 4.2. In most comparisons, the witnesses associated with the focus concepts outnumber those associated with the supporting concepts in each of the three refsets. This implies that the definitions of the focus concepts have undergone a considerable number of modifications during the course of the version releases.

When comparing the January and July 2017 versions of the ICNP Diagnosis refset in UI-Diff(1707, 1701), a considerably higher number (109) of witnesses surfaced, indicating lost information. Seven of these witnesses are linked to supporting concepts; one is a stated witness, while the others are inferred. On the other hand, there are fewer witnesses (25), which represent obtained information that was entailed by July 2017 but was no longer entailed by January 2017. The majority of these witnesses (23) were inferred rather than stated in the ICNP-Diagnosis Subontology-1707. These inferred witnesses point to concealed modifications that the user may wish to investigate.

The ICNP-Intervention changes are quite modest, and the comparison UI-Diff(1707, 1701) appears to have only one inferred focus concept witness. Such a witness may be worth investigating because it may reveal an unanticipated consequence as a result of changes to the ontology.

In the following, we will examine some examples from the evaluation results to demonstrate cases where our approach has successfully generated relevant differences to the chosen focus symbols.

Case (1). Among the focus set witnesses in the GPFP comparison (1707, 1701), the focus concept ***Vocal cord palsy*** has three witnesses, two of which are inferred (in $\mathcal{IW}_{\Sigma_F}$), and one of them is stated (in $\mathcal{SW}_{\Sigma_F}$). The inferred witnesses are the following:

$$\textbf{\textit{Vocal cord palsy}} \sqsubseteq \textit{Disorder of respiratory system} \tag{1}$$

$$\textbf{\textit{Vocal cord palsy}} \sqsubseteq \textit{Paralytic syndrome} \tag{2}$$

Inferred witnesses are entailments that are not stated explicitly in the older ontology. The reason for the inferred witnesses can be determined using an OWL justification tool [12] or through manual inspection. We demonstrate the causes of inferred witnesses using the OWL justification tool[3] on the inferred witness set and the subontology-1701. The axioms that explain the inferred witness (1) are as follows:

$$\begin{aligned}&\textbf{\textit{Vocal cord palsy}} \sqsubseteq \textit{Paralysis of vocal cords or larynx}\\&\quad \sqcap \exists \textit{Role group.}(\exists \textit{Finding site.Structure of nervous system})\\&\quad \sqcap \exists \textit{Role group.}(\exists \textit{Finding site.Vocal cord structure})\end{aligned} \tag{3}$$

$$\begin{aligned}&\textit{Paralysis of vocal cords or larynx}\\&\quad \sqsubseteq \textit{Disorder of respiratory system}\end{aligned} \tag{4}$$

We can observe that ***Vocal cord palsy*** is subsumed by *Disorder of respiratory system* via the concept *Paralysis of vocal cords or larynx*. This subsumption relationship was revealed by forgetting the concept *Paralysis of vocal cords or larynx*, which resulted in the aforementioned witness (1).

Upon generating the explanation axioms for the inferred witnesses, several actions can be done by the modeler, including logging the witness set along with their explanation axioms for future improvements of the ontology, as well as understanding the effect of the change made on the ontology.

One stated witness associated with the focus concept ***Vocal cord palsy*** was found in $\mathcal{SW}_{\Sigma_F}$, which is:

$$\begin{aligned}&\textbf{\textit{Vocal cord palsy}} \sqsubseteq\\&\quad \exists \textit{Role group.}(\exists \text{Finding site.Structure of nervous system}).\end{aligned} \tag{5}$$

[3]https://github.com/matthewhorridge/owlexplanation

It is worth noting that the stated witness of axiom (5) for ***Vocal cord palsy*** is a result of abstracting the focus concept definition. Abstracting the definition of ***Vocal cord palsy*** aids in understanding which necessary condition that ***Vocal cord palsy*** inherits from *Paralysis of vocal cords or larynx* that is not entailed by the subontology version 1707. In contrast, computing the semantic difference between the original ontologies rather than the subontologies does not show the witness axiom (5), which may result in insufficient comprehension of all the modifications to the definition of the focus concept ***Vocal cord palsy***.

Case (2). SNOMED CT modelers routinely perform version control tasks such as reviewing changes to the concept hierarchy [2]. When dealing with patient case queries, such changes become crucial since they may alter the subsumption relationships rather than the actual patient cases, resulting in inconsistencies in reporting trends [20, 21]. According to a SNOMED CT terminologist [2], there are times when modelers become particularly interested to learn about changes to the hierarchy of supporting concepts especially when supporting concepts appear in definitions of concepts that remain unaltered between versions. We present an example from the evaluation results to show such cases.

Example. Looking at the set of stated supporting concept witnesses ($\mathcal{SW}_{\Sigma_S}$) in the comparison ICNP-Intervention(1707, 1701), there is only one stated witness associated with the supporting concept *Finding related to ability to walk*. The witness is:

$$\begin{aligned}&\underline{\textit{Finding related to ability to walk}}\\&\quad \sqsubseteq \textit{Finding of activity of daily living}\end{aligned} \tag{6}$$

Finding related to ability to walk is a supporting concept of the focus concept ***Assessment of ability to walk*** definition:

$$\begin{aligned}&\textbf{\textit{Assessment of ability to walk}} \sqsubseteq \textit{Procedure} \sqcap\\&\quad \exists \textit{RoleGroup.}(\exists \textit{Method.Evaluation - action}) \sqcap\\&\quad \exists \textit{RoleGroup.}(\exists \textit{Has focus.}\underline{\textit{Finding related to ability to walk}})\end{aligned} \tag{7}$$

The definition of the focus concept ***Assessment of ability to walk*** is identically defined in both versions of the subontologies 1701 and 1707. The witness (6) indicates that there was a change in the position of the supporting concept *Finding related to ability to walk* in the supporting concept hierarchy. This witness is entailed by the subontology version 1701 but not by 1707. To understand the new position of the supporting concept *Finding related to ability to walk*, we can check the newer subontology version 1707 and find that it has been changed to be subsumed by *Clinical finding*. We also remark that computing the UI-Diff between the ⊥-modules [29] does not reveal witness (6) associated with the supporting concept *Finding related to ability to walk*.[4] This demonstrates that computing the UI-Diff on focus set subontologies rather than the original ontologies reveals changes associated with supporting concepts that are not revealed using other approaches.

[4]As we did not have access to the results of the UI-Diff comparisons between the original ontologies, we performed a spot check that involved computing the bottom modules for Σ_F. Then, we computed the UI-Diff between the produced ⊥-modules to see if computing the UI-Diff between the original ontologies (represented by the ⊥-modules) would reveal the witness connected with the supporting concept (*Finding related to ability to walk*). Our test yielded no such witness.

The example shows that subontologies play a crucial role in detecting specific modifications associated with supporting concepts that are related to the definitions of the focus concepts. This is because the information contained inside subontologies is limited to the definitions of concepts of interest (the focus concepts) and the hierarchical relationships of their supporting concepts.

6 CONCLUSION

We introduced a method for tracking semantic differences between ontologies based on novel combinations of ontology extraction methods. This is accomplished by first extracting subontologies for a given set of focus symbols, which are then utilized to compute semantic differences. This leads in a focused set of semantic changes that are pertinent to selected subdomains of the content of the ontologies. We illustrated that focus set subontontologies aid in the identification of additional witnesses linked with focus concepts. These additional witnesses alert modelers to potential changes in concept meaning caused by the addition or removal of defining characteristics from focus concept definitions in released revisions. The segmentation of witnesses highlights the utility of our method in simplifying analysis tasks when the modeler is primarily concerned with discovering distinctions between focus and supporting concepts. Furthermore, distinguishing the difference between stated and inferred witnesses helps the modeler capture unobserved inferred witnesses for additional analysis. This shows the method's usefulness in locating potentially obscure witnesses associated with a specific refset or other concept sets.

REFERENCES

[1] Ghadah Alghamdi. 2022. *Focus Set Subontologies and Semantic Differences for Large $\mathcal{ELH}$ Ontologies.* Ph.D. Dissertation. University of Manchester, UK.

[2] Ghadah Alghamdi, Yongsheng Gao, and Renate A Schmidt. [n.d.]. Focus Set Semantic Differences Meeting on 14-12-2021. Meeting to discuss the approach in Section 4.1 and results outlined in Section 5.2 with SNOMED CT modelers.

[3] Ghadah Alghamdi, Renate A. Schmidt, Warren Del-Pinto, and Yongsheng Gao. 2021. Upwardly Abstracted Definition-Based Subontologies. In *K-CAP '21: Knowledge Capture Conference.* ACM, 209–216.

[4] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. 2017. *An Introduction to Description Logic.* Cambridge University Press.

[5] Jieying Chen, Ghadah Alghamdi, Renate A. Schmidt, Dirk Walther, and Yongsheng Gao. 2019. Ontology Extraction for Large Ontologies via Modularity and Forgetting. In *Proceedings of the 10th International Conference on Knowledge Capture, K-CAP.* ACM, 45–52.

[6] Silvio Ghilardi, Carsten Lutz, and Frank Wolter. 2006. Did I Damage My Ontology? A Case for Conservative Extensions in Description Logics. In *Proc. KR'06.* AAAI Press, 187–197.

[7] Jennifer Golbeck, Gilberto Fragoso, Frank W. Hartel, James A. Hendler, Jim Oberthaler, and Bijan Parsia. 2003. The National Cancer Institute's Thésaurus and Ontology. *J. Web Semant.* 1 (2003), 75–80.

[8] Rafael S. Gonçalves, Bijan Parsia, and Ulrike Sattler. 2012. Concept-Based Semantic Difference in Expressive Description Logics. In *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference (Lecture Notes in Computer Science, Vol. 7649).* Springer, 99–115.

[9] Rafael S. Gonçalves, Bijan Parsia, and Ulrike Sattler. 2012. Ecco: A Hybrid Diff Tool for OWL 2 Ontologies. In *Proceedings of OWL: Experiences and Directions Workshop 2012 (CEUR Workshop, Vol. 849).* CEUR-WS.org.

[10] Rinke Hoekstra, Joost Breuker, Marcello Di Bello, and Alexander Boer. 2009. LKIF Core: Principled Ontology Development for the Legal Domain. In *Law, Ontologies and the Semantic Web - Channelling the Legal Information Flood (Frontiers in Artificial Intelligence and Applications, Vol. 188).* IOS Press, 21–52.

[11] Matthew Horridge and Sean Bechhofer. 2009. The OWL API: A Java API for Working with OWL 2 Ontologies, In Proceedings of the 6th International Conference on OWL: Experiences and Directions. *Semantic Web* 2, 11–21.

[12] Matthew Horridge, Bijan Parsia, and Ulrike Sattler. 2015. *The OWL Explanation Workbench: A Toolkit for Working with Justifications for Entailments in OWL Ontologies.* Technical Report. The Semantic Web Journal.

[13] SNOMED International. [n.d.]. SNOMED CT - Practical Guide to Reference Sets: https://confluence.ihtsdotools.org/display/DOCRFSPG. Accessed: 2022-02-28.

[14] SNOMED International. [n.d.]. SNOMED CT Release Types: https://confluence.ihtsdotools.org/display/DOCRELFMT/3.2+Release+Types. Accessed: 2023-06-19.

[15] Ernesto Jiménez-Ruiz, Bernardo Cuenca Grau, Ian Horrocks, and Rafael Berlanga Llavori. 2011. Supporting concurrent ontology development: Framework, algorithms and tool. *Data Knowl. Eng.* 70, 1 (2011), 146–164.

[16] Aditya Kalyanpur, Bijan Parsia, Matthew Horridge, and Evren Sirin. 2007. Finding All Justifications of OWL DL Entailments. In *The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007 (Lecture Notes in Computer Science, Vol. 4825).* Springer, 267–280.

[17] Boris Konev, Carsten Lutz, Dirk Walther, and Frank Wolter. 2008. Logical Difference and Module Extraction with CEX and MEX. In *Proceedings of the 21st International Workshop on Description Logics (DL2008) (CEUR Workshop, Vol. 353).*

[18] Boris Konev, Dirk Walther, and Frank Wolter. 2008. The Logical Difference Problem for Description Logic Terminologies. In *Automated Reasoning, 4th International Joint Conference, IJCAR 2008 (Lecture Notes in Computer Science, Vol. 5195).* Springer, 259–274. https://doi.org/10.1007/978-3-540-71070-7_21

[19] Boris Konev, Dirk Walther, and Frank Wolter. 2009. Forgetting and Uniform Interpolation in Large-Scale Description Logic Terminologies. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI).* 830–835.

[20] Dennis Lee, Ronald Cornet, and Francis Lau. 2011. Implications of SNOMED CT Versioning. *International Journal of Medical Informatics* 80, 6 (2011), 442–453.

[21] Dennis Lee, Nicolette F. de Keizer, Francis Y. Lau, and Ronald Cornet. 2014. Literature Review of SNOMED CT Use. *Journal of the American Medical Informatics Association : JAMIA* 21 e1 (2014), e11–9.

[22] Zhao Liu, Chang Lu, Ghadah Alghamdi, Renate A. Schmidt, and Yizheng Zhao. 2021. Tracking Semantic Evolutionary Changes in Large-Scale Ontological Knowledge Bases. In *CIKM'21: The 30th ACM International Conference on Information and Knowledge Management.* ACM, 1130–1139.

[23] Michel Ludwig and Boris Konev. 2014. Practical Uniform Interpolation and Forgetting for $\mathcal{ALC}$ TBoxes with Applications to Logical Difference. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference.* AAAI Press.

[24] Carsten Lutz and Frank Wolter. 2010. Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. *J. Symb. Comput.* 45, 2 (2010), 194–228.

[25] Carsten Lutz and Frank Wolter. 2011. Foundations for Uniform Interpolation and Forgetting in Expressive Description Logics. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence.* IJCAI/AAAI, 989–995.

[26] Nadeschda Nikitina and Sebastian Rudolph. 2014. (Non-)Succinctness of uniform interpolants of general terminologies in the description logic $\mathcal{EL}$. *Artif. Intell.* 215 (2014), 120–140.

[27] Christopher Ochs, James T. Case, and Yehoshua Perl. 2016. Tracking the Remodeling of SNOMED CT's Bacterial Infectious Diseases. In *AMIA 2016, American Medical Informatics Association Annual Symposium.* AMIA.

[28] Christopher Ochs, Yehoshua Perl, Gai Elhanan, and James T. Case. 2016. A Descriptive Delta for Identifying Changes in SNOMED CT. In *Proceedings of the Joint International Conference on Biological Ontology and BioCreative (CEUR Workshop, Vol. 1747).*

[29] Ulrike Sattler, Thomas Schneider, and Michael Zakharyaschev. 2009. Which Kind of Module Should I Extract?. In *Proceedings of the 22nd International Workshop on Description Logics (DL 2009) (CEUR Workshop, Vol. 477).*

[30] Miguel-Ángel Sicilia, Elena García-Barriocanal, Salvador Sánchez-Alonso, and Daniel Rodríguez-García. 2009. Ontologies of engineering knowledge: general structure and the case of software engineering. *The Knowledge Engineering Review* 24, 3 (2009), 309–326.

[31] Alessandro Solimando, Ernesto Jiménez-Ruiz, and Giovanna Guerrini. 2017. Minimizing conservativity violations in ontology alignments: algorithms and evaluation. *Knowledge and Information Systems* 51, 3 (2017), 775–819.

[32] Kent A. Spackman. 2000. SNOMED RT and SNOMED CT. Promise of an international clinical ontology. *M.D. Computing* 17 (2000), 29.

[33] Kent A. Spackman. 2001. Normal forms for description logic expressions of clinical concepts in SNOMED RT. In *American Medical Informatics Association Annual Symposium.* AMIA.

[34] The Gene Ontology Consortium. 2018. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research* 49 (2018), gky1055.

[35] Chiara Del Vescovo, Pavel Klinov, Bijan Parsia, Ulrike Sattler, Thomas Schneider, and Dmitry Tsarkov. 2013. Empirical Study of Logic-Based Modules: Cheap Is Cheerful. In *Proceedings of the Semantic Web - ISWC 2013*, Vol. 8218. Springer, 84–100.

[36] W3C OWL Working Group. 2012. OWL 2 Web Ontology Language Document Overview (Second Edition) - W3C Recommendation 11 Dec 2012. http://www.w3.org/TR/owl2-overview/

[37] Yizheng Zhao, Ghadah Alghamdi, Renate A. Schmidt, Hao Feng, Giorgos Stoilos, Damir Juric, and Mohammad Khodadadi. 2019. Tracking Logical Difference in Large-Scale Ontologies: A Forgetting-Based Approach. In *Proc. of AAAI'19.* AAAI Press, 3116–3124.