Prevalence and Effects of Class Hierarchy Precompilation in Biomedical Ontologies

Christian Kindermann(✉), Bijan Parsia, and Uli Sattler

University of Manchester, Manchester M13 9PL, UK
{christian.kindermann,bijan.parsia,uli.sattler}@manchester.ac.uk

Abstract. It is sometimes claimed that adding inferred axioms, e.g. the inferred class hierarchy (ICH), to an ontology can improve reasoning performance or an ontology's usability in practice. While such beliefs may have an effect on how ontologies are published, there is no conclusive empirical evidence to support them. To develop an understanding of the impact of this practice, both for ontology curators as well as tools, we survey to what extent published ontologies in BioPortal already contain their ICH and most specific class assertions (MSCA). Furthermore, we investigate how added inferred axioms from these sets can affect the performance of standard reasoning tasks such as classification and realisation. We find that axioms from the ICH and MSCA are highly prevalent in published biomedical ontologies. Our reasoning evaluation indicates that added inferred axioms are likely to be inconsequential for reasoning performance. However, we observe instances of both positive as well as negative effects that seem to depend on the used reasoner for a given ontology. These results suggest that the practice of adding inferred axioms during the release process of ontologies should be subject to a task-specific analysis that determines whether desired effects are obtained.

Keywords: Ontology engineering · Reasoning performance · OWL · Web ontology language · BioPortal · Class hierarchy · Concept hierarchy

1 Introduction

In the biomedical domain, there seems to be a belief that adding certain kinds of inferred axioms to an ontology, e.g., its inferred class hierarchy (ICH), may improve the ontology's usability in practice. This is even said to be a fundamental step in the release process of ontologies and is supported by automation tools [7]. However, there are also arguments claiming that redundant subsumption axioms can negatively affect the maintenance burden for ontology curators [13]. Overall, there appears to be a lot of folk-wisdom about the benefits and drawbacks of materialising entailed axioms.

To develop an understanding of the possible impact of this practice, we introduce the notion of *precompilation* to distinguish between substantive and

J. Z. Pan et al. (Eds.): ISWC 2020, LNCS 12506, pp. 330–348, 2020.
https://doi.org/10.1007/978-3-030-62419-4_19

redundant materialisations of entailment sets. We survey to what extent published biomedical ontologies materialise entailment sets derived from the ICH, and how this practice affects reasoning performance. Our results indicate that axioms of such sets are often materialised. While many ontologies contain redundant axioms from the ICH, their relative proportion is often low. We find that precompiling the ICH can positively impact the performance of reasoning tasks. However, this does not hold in general and depends on the reasoning task, the reasoner, and the given ontology itself.

2 Preliminaries

We assume the reader to be familiar with OWL, in particular OWL 2 [3], and only fix some terminology. Let N_C, N_I, and N_P be sets of *class names*, *individual names*, and *property names*. A *class* is either a class name or a *complex class* built using OWL class constructors. In the following, we use DL notation for increased readability; in particular, we use $\mathsf{A} \sqsubseteq \mathsf{B}$ for a subclass axiom between A and B, $\mathsf{A} \equiv \mathsf{B}$ for an equivalence axiom between A and B, $\bot, \top$ for `owl:Thing`, `owl:Nothing`, $\mathsf{A}(a)$ for a class assertion between an individual a and a class A, and write $\mathcal{O} \models \alpha$ to denote that the ontology $\mathcal{O}$ entails the axiom α. We also use $\equiv(\mathsf{A}_1, \ldots, \mathsf{A}_n)$ to denote the n-ary equivalence axiom between the classes A_i and take the OWL view that its parameters are a set, i.e, $\equiv(\mathsf{A}_1, \mathsf{A}_2) = \equiv(\mathsf{A}_2, \mathsf{A}_1)$. In particular, we say that the axioms $\mathsf{A} \sqsubseteq \mathsf{B}$, $\equiv(\mathsf{A}_1, \ldots, \mathsf{A}_n)$ are *atomic* if $\mathsf{A}, \mathsf{B}, \mathsf{A}_1, \ldots, \mathsf{A}_n \in N_C \cup \{\bot, \top\}$. Similarly, we say that the assertion $\mathsf{A}(a)$ is *atomic* if $\mathsf{A} \in N_C \cup \{\bot, \top\}$. Other axioms are called *complex*.

Furthermore, we use $[\mathsf{A}]$ to denote the set $\{\mathsf{A}_i \mid \mathcal{O} \models \mathsf{A} \equiv \mathsf{A}_i\}$ for a class name A in ontology $\mathcal{O}$. By abuse of notation we write $[\mathsf{A}] \sqsubseteq [\mathsf{B}]$ to denote the set of axioms $\{\mathsf{A}' \sqsubseteq \mathsf{B}' \mid \mathsf{A}' \in [\mathsf{A}], \mathsf{B}' \in [\mathsf{B}]\}$ and $[\mathsf{A}](a)$ to denote the set $\{\mathsf{A}_i(a) \mid \mathsf{A} \equiv \mathsf{A}_i\}$.

An *ontology* is a set of axioms. An ontology is *logically empty* if it entails only tautologies. We write $\mathcal{O}_1 \equiv \mathcal{O}_2$ to denote that $\mathcal{O}_1$ and $\mathcal{O}_2$ are equivalent, i.e., have the same models, and use $\mathcal{O}$ for an ontology and $\widetilde{\mathcal{O}}$ for the set of class, property, and individual names in $\mathcal{O}$. Finally, since we are concerned with entailments, we only consider consistent ontologies.

3 Precompilation

We introduce the notion of *precompilation* to capture the idea of systematically adding inferred axioms to an ontology. The characteristics of related axioms are defined in terms of *entailment sets*.

Definition 1 (Entailment Set, Materialisation). *Let $\mathcal{O}$ be an ontology. An* entailment set *of $\mathcal{O}$ is a set of axioms $\mathcal{E}$ such that $\mathcal{O} \models \alpha$ for each $\alpha \in \mathcal{E}$. An entailment set $\mathcal{E}$ is* materialised *in $\mathcal{O}$ if $\mathcal{E} \subseteq \mathcal{O}$.*

Note that an axiom in a materialised entailment set is not necessarily entailed by the remainder of the ontology. For example, consider an ontology $\mathcal{O}$ with a single axiom α. Clearly, the set $\{\alpha\}$ is an entailment set of $\mathcal{O}$ and is materialised. However, unless α is a tautology, removing it from $\mathcal{O}$ changes $\mathcal{O}$'s meaning. Hence, a materialisation may in fact be a substantive part of the ontology as opposed to a semantically redundant addition.

Definition 2 (Redundancy). *An entailment set $\mathcal{E}$ of an ontology $\mathcal{O}$ is* redundant in $\mathcal{O}$ *if $\mathcal{E}$ is also an entailment set of $\mathcal{O} \setminus \mathcal{E}$.*

In the following, we often call an axiom α redundant in $\mathcal{O}$ as a shorthand for $\{\alpha\}$ being redundant in $\mathcal{O}$. Of course, adding entailed axioms to an ontology adds redundancy. To capture the idea of purposefully adding sets of entailed axioms as a form of preprocessing, we propose the notion of *precompilation* as redundant materialisations.

Definition 3 (Precompilation). *An entailment set $\mathcal{E}$ of an ontology $\mathcal{O}$ is* precompiled in $\mathcal{O}$ *if $\mathcal{E}$ is materialised and redundant in $\mathcal{O}$.*

Any entailment set can be partitioned into three (possibly empty) subsets of precompiled, materialised but not redundant, and non-materialised axioms. Note that such a partition is not necessarily unique as we will explain in Sect. 4.

A standard OWL reasoning service is *classification*, i.e., the computation of the entailment set of all atomic subsumption axioms. As discussed in [1], fixing reasonable sets of even atomic entailments is tricky.

Definition 4 (Inferred Class Hierarchy). *The* inferred class hierarchy $\mathsf{ICH}(\mathcal{O})$ *of $\mathcal{O}$ is defined as follows:*

$$\begin{aligned}\mathsf{ICH}(\mathcal{O}) = \; & \{\mathsf{A} \sqsubseteq \mathsf{B} \mid \mathsf{A}, \mathsf{B} \in N_C \cup \{\bot, \top\}, \mathcal{O} \models \mathsf{A} \sqsubseteq \mathsf{B}\} \cup \\ & \{\equiv(\mathsf{A}_1, \ldots, \mathsf{A}_n) \mid \mathsf{A}_1, \ldots, \mathsf{A}_n \in N_C \cup \{\bot, \top\}, \mathcal{O} \models \equiv(\mathsf{A}_1, \ldots, \mathsf{A}_n)\}.\end{aligned}$$

While the inferred class hierarchy of an ontology is a well understood and widely used (finite) entailment set, it has been noted that informal references to this set are often understood as some more restrictive subset [1]. In our case, we include redundant versions of equivalence axioms, e.g., if $\equiv(\mathsf{A}_1, \mathsf{A}_2, \mathsf{A}_3) \in \mathsf{ICH}(\mathcal{O})$, then $\equiv(\mathsf{A}_1, \mathsf{A}_2) \in \mathsf{ICH}(\mathcal{O})$. In practice, the ICH is most commonly represented by some form of a transitive *reduct* and may include or exclude tautologies, e.g., $\mathsf{A} \sqsubseteq \top$ or $\bot \sqsubseteq \mathsf{A}$. Therefore, we distinguish between four distinct entailment sets that capture different aspects of an ontology's ICH.

Definition 5 (Transitive Reduct). *A* transitive reduct *of* $\mathsf{ICH}(\mathcal{O})$*, written* $\mathsf{TR}(\mathcal{O})$*, is*

1. *a subset of* $\mathsf{ICH}(\mathcal{O})$ *that is equivalent to* $\mathsf{ICH}(\mathcal{O})$ *and*
2. *cardinality minimal, i.e., if $\mathcal{O}' \subseteq \mathsf{ICH}(\mathcal{O})$ and $\mathcal{O}' \equiv \mathcal{O}$, then $|\mathsf{TR}(\mathcal{O})| \leq |\mathcal{O}'|$.*

Note that a transitive reduct is not necessarily unique in the presence of equivalences. Consider $\mathcal{O}_{ex} = \{\mathsf{AlaskanMoose} \sqsubseteq \mathsf{Moose}, \mathsf{Moose} \equiv \mathsf{Elk}, \mathsf{Elk} \sqsubseteq \mathsf{Deer}\}$ as an example. Here, $\mathsf{TR}(\mathcal{O}_{ex}) = \mathcal{O}_{ex}$ is only one of four transitive reducts. Also, the transitive reduct mentions $\bot$ iff the ontology contains unsatisfiable classes (which are all gathered in a single maximal equivalence axiom that includes $\bot$). Dually, it mentions $\top$ iff the ontology contains global classes (which are all gathered in a single maximal equivalence axiom that includes $\top$). In case there is only a single global class, this leads to two reducts, one with an equivalence class and one with a subclass axiom with/of $\top$.

Second, we define the set of tautologies that would be included in the transitive reduction of an ontology's class hierarchy if $\bot$ and $\top$ were "normal" class names.

Definition 6 (Tautological Completion). *The* tautological completion $\top\bot\text{-}\mathsf{TR}(\mathcal{O})$ *of* $\mathcal{O}$ *is defined as follows:*

$$\begin{aligned}\top\bot\text{-}\mathsf{TR}(\mathcal{O}) = \; & \{\mathsf{A} \sqsubseteq \top \mid \mathsf{A} \in N_C \setminus [\top] \textit{ and } \mathcal{O} \models \mathsf{A} \sqsubseteq \mathsf{B} \textit{ implies } \mathsf{B} \in [\mathsf{A}] \cup [\top]\} \cup \\ & \{\bot \sqsubseteq \mathsf{A} \mid \mathsf{A} \in N_C \setminus [\bot] \textit{ and } \mathcal{O} \models \mathsf{B} \sqsubseteq \mathsf{A} \textit{ implies } \mathsf{B} \in [\mathsf{A}] \cup [\bot]\}.\end{aligned}$$

In this definition, $\mathsf{TR}(\mathcal{O})$ occurs unquantified as it does not matter which one we pick in case there are more than one: they only differ in subclass axioms to and from equivalent classes and do not contain tautologies, hence $\top\bot\text{-}\mathsf{TR}(\mathcal{O})$ always contains all subclass axioms between top-level classes (not equivalent to $\top$) and $\top$ and between $\bot$ and bottom-level classes (not equivalent to $\bot$). Continuing our example $\mathcal{O}_{ex}$, we have $\top\bot\text{-}\mathsf{TR}(\mathcal{O}_{ex}) = \{\bot \sqsubseteq \mathsf{AlaskanMoose}, \mathsf{Deer} \sqsubseteq \top\}$.

Third, we define short-cuts in the class hierarchy, i.e., non-tautological but inferred subsumption axioms that are not in any transitive reduction of an ontology's class hierarchy.

Definition 7 (Short Cut). *The set of* short cuts $\mathsf{SC}(\mathcal{O})$ *is an defined as follows:*

$$\mathsf{SC}(\mathcal{O}) = \mathsf{ICH}(\mathcal{O}) \setminus \bigcup_{\mathsf{TR}(\mathcal{O})} (\mathsf{TR}(\mathcal{O}) \cup \top\bot\text{-}\mathsf{TR}(\mathcal{O})).$$

Please note that short cuts can also contain equivalence axioms. For example, consider $\mathcal{O}' = \mathcal{O}_{ex} \cup \{\mathsf{Elk} \equiv \mathsf{AlcesAlces}\}$. Then we have $\mathsf{Moose} \equiv \mathsf{Elk}, \mathsf{Elk} \equiv \mathsf{AlcesAlces} \in \mathsf{SC}(\mathcal{O}')$ because $\equiv(\mathsf{Moose}, \mathsf{Elk}, \mathsf{AlcesAlces}) \in \mathsf{TR}(\mathcal{O}')$ (for all transitive reducts of $\mathcal{O}'$). Also note that short cut axioms are not necessarily redundant in an ontology as demonstrated by the example. Lastly, we define short cut tautologies in an ontology's class hierarchy that are not in the tautological completion of its transitive reduction.

Definition 8 (Short Cut Tautologies). *The set of* short cut tautologies, *written* $\top\bot\text{-}\mathsf{SC}(\mathcal{O})$, *is defined as follows:*

$$\begin{aligned}\top\bot\text{-}\mathsf{SC}(\mathcal{O}) = \; & \{\mathsf{A} \sqsubseteq \top \mid \mathsf{A} \in \widetilde{\mathcal{O}}, \mathsf{A} \sqsubseteq \top \not\in \top\bot\text{-}\mathsf{TR}(\mathcal{O})\} \cup \\ & \{\bot \sqsubseteq \mathsf{A} \mid \mathsf{A} \in \widetilde{\mathcal{O}}, \bot \sqsubseteq \mathsf{A} \not\in \top\bot\text{-}\mathsf{TR}(\mathcal{O})\} \cup \\ & \{\mathsf{A} \sqsubseteq \mathsf{A} \mid \mathsf{A} \in N_C \cup \{\bot, \top\}\}.\end{aligned}$$

In the case of our example $\mathcal{O}_{ex}$, we have for instance $\mathsf{AlaskanMoose} \sqsubseteq \top \in \top\bot\text{-}\mathsf{SC}(\mathcal{O}_{ex})$. The distinction between atomic axioms in (i) transitive reducts, (ii) the tautological completion of transitive reducts, (iii) non-tautological short cuts, and (iv) short cut tautologies results in a unique partition of the $\mathsf{ICH}(\mathcal{O})$ for a given transitive reduct.

In addition to $\mathsf{ICH}(\mathcal{O})$, which only captures the terminological knowledge about named class, another important entailment set in practice is the set of *class assertions* for individuals contained in an ontology.

Definition 9 (Inferred Class Hierarchy Assertions). *The set of* inferred class hierarchy assertions *of* $\mathcal{O}$ *is defined as follows:*

$$\mathsf{ICHA}(\mathcal{O}) = \{\mathsf{A}(a) \mid \mathsf{A} \in \widetilde{\mathcal{O}}, a \in N_I, \mathcal{O} \models \mathsf{A}(a)\}$$

As for $\mathsf{ICH}(\mathcal{O})$, there may be some variance in terms of how $\mathsf{ICHA}(\mathcal{O})$ is understood. Therefore we define the set of *most specific* class assertions as the smallest entailment set that still captures the $\mathsf{ICHA}(\mathcal{O})$, which is realised as another standard OWL reasoning service called *realisation.*

Definition 10 (Most Specific Class Assertions). *A set of* most specific class assertions *of* $\mathcal{O}$*, written* $\mathsf{MSCA}(\mathcal{O})$*, is a (cardinality) minimal set* $\mathsf{MSCA}(\mathcal{O}) \subseteq \mathsf{ICHA}(\mathcal{O})$ *such that* $\mathsf{MSCA}(\mathcal{O}) \cup \mathsf{ICH}(\mathcal{O}) \models \mathsf{ICHA}(\mathcal{O})$.

Extending $\mathcal{O}_{ex}$ with $\{\mathsf{Elk}(a), \mathsf{Deer}(a), \mathsf{hasCalf}(a, b)\}$, yields $\mathsf{MSCA}(\mathcal{O}_{ex}) = \{\mathsf{Elk}(a)\}$.

Analogously to what we did for the ICH, we define the (unique) tautological completion for MSCAs:

Definition 11 (Tautological Completion). *The* tautological completion *of* $\mathsf{MSCA}(\mathcal{O})$ *is defined as follows:*

$$\top\bot\text{-}\mathsf{MSCA}(\mathcal{O}) = \{\top(a) \mid a \in \widetilde{\mathcal{O}} \textit{ and there is no } \mathsf{A}(a) \in \mathsf{MSCA}(\mathcal{O})\}.$$

In case of our extended example $\mathcal{O}_{ex}$, we have $\top\bot\text{-}\mathsf{MSCA}(\mathcal{O}_{ex}) = \{\top(b)\}$.

Similarly, we define a notion for short-cuts w.r.t. class assertions:

Definition 12 (Short Cut Assertions). *The set of* short cut assertions *of* $\mathsf{SCA}(\mathcal{O})$ *is defined as follows:*

$$\mathsf{SCA}(\mathcal{O}) = \mathsf{ICHA}(\mathcal{O}) \setminus \bigcup_{\mathsf{MSCA}(\mathcal{O})} (\mathsf{MSCA}(\mathcal{O}) \cup \top\bot\text{-}\mathsf{MSCA}(\mathcal{O})).$$

Continuing our extended example $\mathcal{O}_{ex}$, we find $\mathsf{Deer}(a) \in \mathsf{SCA}(\mathcal{O}_{ex})$.

And finally, we define a notion for short-cut assertion tautologies:

Definition 13 (Short Cut Assertion Tautologies). *The set of* short cut assertion tautologies $\top\bot\text{-}\mathsf{SCA}(\mathcal{O})$ *is defined as follows:*

$$\top\bot\text{-}\mathsf{SCA}(\mathcal{O}) = \{\top(a) \mid \mathcal{O} \models \top(a)\} \setminus \top\bot\text{-}\mathsf{MSCA}(\mathcal{O}).$$

Analogously to the case of the ICH, the Definitions 10–13 give rise to a unique partition of the ICHA into four sets for a given set of most specific class assertions.

4 Determining the Extent of Precompilation

In this section, we discuss how we can determine the extent of precompilation in an ontology. Given an entailment set $\mathcal{E}$ of an ontology $\mathcal{O}$, we check whether it is materialised in $\mathcal{O}$ by simply checking whether $\mathcal{E} \subseteq \mathcal{O}$ holds. To determine whether $\mathcal{E}$ is precompiled, we simply test whether $\mathcal{O} \setminus \mathcal{E} \models \mathcal{E}$ holds. This straightforward way of determining precompilation suffers, however, from two issues: firstly, it is insensitive to different but equivalent representations of an entailment set. For example, a subsumption entailed from an equivalence axiom is not necessarily a precompiled axiom. Secondly, it considers the entailment set as a whole: if a single axiom of a large entailment set is not precompiled, then the whole entailment set is not precompiled. Therefore, instead of searching for precompiled entailment sets, it is more appropriate to search for maximal precompiled subsets of a given entailment set in an ontology.

Identifying sets of redundant axioms in ontologies is known to be challenging in practice [6,13]. While it is straightforward to identify a single axiom α in an ontology $\mathcal{O}$ as redundant by testing whether $\mathcal{O} \setminus \{\alpha\} \models \alpha$ holds, such axioms do not, in general, form redundant subsets when grouped together. As an example, consider the ontology $\mathcal{O} = \{\mathsf{A} \sqsubseteq \mathsf{B}, \mathsf{A} \sqsubseteq \mathsf{C}, \mathsf{A} \sqsubseteq \mathsf{B} \sqcap \mathsf{C}\}$. Then for all axioms $\alpha \in \mathcal{O}$, we have $\mathcal{O} \setminus \{\alpha\} \models \alpha$. However, all three axioms taken together, i.e. $\mathcal{O}$ itself, does not constitute a redundant set. As a consequence of this, removing redundant axioms from an ontology comes down to a choice between a number of alternatives. This also means, that for a given ontology, there may exist several irredundant equivalent ontologies.

Since we are interested in identifying redundant axioms with respect to some entailment set, we define a notion of irredundancy for ontologies accordingly.

Definition 14 (Reduced Ontology). *Let $\mathcal{O}$ be an ontology and $\mathcal{E}$ an entailment set of $\mathcal{O}$. An ontology $\mathcal{O}^- \subseteq \mathcal{O}$ is a* reduction of $\mathcal{O}$ with respect to $\mathcal{E}$*, if*

(i) $\mathcal{O} \setminus \mathcal{O}^- \subseteq \mathcal{E}$,
(ii) $\mathcal{O}^- \equiv \mathcal{O}$,
(iii) there exists no $\alpha \in \mathcal{E}$ such that $\{\alpha\}$ is redundant in $\mathcal{O}^-$.

Each reduction $\mathcal{O}^-$ of $\mathcal{O}$ can be associated with its corresponding precompiled subset of $\mathcal{E}$, namely $\mathcal{O} \setminus \mathcal{O}^-$. With this, we can elaborate on the statement made in Sect. 3 with respect to possible partitions of an entailment set into subsets of precompiled, materialised but not redundant, and non-materialised axioms.

Proposition 1. *A reduction $\mathcal{O}^-$ of an ontology $\mathcal{O}$ wrt. an entailment set $\mathcal{E}$ uniquely identifies a partition of $\mathcal{E}$ into three subsets defined as follows:*

1. *$P = \mathcal{O} \setminus \mathcal{O}^-$ precompiled axioms in $\mathcal{O}$,*
2. *$M = \mathcal{O}^- \cap \mathcal{E} = (\mathcal{O} \cap \mathcal{E}) \setminus P$ of materialised but not redundant axioms in $\mathcal{O}^-$,*
3. *$N = \mathcal{E} \setminus \mathcal{O}$ non-materialised axioms in $\mathcal{O}$.*

5 Methods

5.1 Materials

Ontology Corpus. We use a publicly available snapshot of BioPortal from 2017.[1] The data set of ontologies with their imports closure merged in encompasses a total of 438 ontologies. We select ontologies for individual reasoners according to the following criteria: (i) can be processed using the OWL API, (ii) contains logical axioms, (iii) is found to be consistent by a reasoner, and (iv) can be classified by a reasoner within one hour. The last criterion is chosen primarily for practical reasons owing to the large scale of our empirical investigation. This choice is justified by empirical evidence that most ontologies can be classified well within one hour [5].

No ontologies have been excluded based on criterion (i). A total of 13 ontologies were excluded based on criterion (ii). As for criteria (iii) and (iv), note that an ontology is not necessarily deemed consistent or classifiable by all reasoners. In particular, HermiT found three ontologies to be inconsistent, Pellet five, JFact eight, and Konclude seven. Likewise, 352 ontologies of the remaining ontologies could by classified by HermiT, 319 by Pellet, 381 by JFact, and 406 by Konclude. Any statements involving a reasoner will be made w.r.t. these the reasoner's respective ontologies.

In experiments, we distinguish between ontologies that consist of atomic axioms only, axioms expressible in $\mathcal{EL}^{++}$, and rich otherwise. We refer to these three kinds of ontologies as *atomic*, $\mathcal{EL}^{++}$, and *rich* ontologies respectively.

Experimental Environment. Ontologies in this study are processed using the OWL API (version 4.5.13). With the exception of Konclude[2] (version 0.6.2), all reasoning tasks are orchestrated via a reasoner's OWL API support. The used reasoners are HermiT[3] (version 1.3.8.413), JFact[4] (version 4.0.4), and Pellet[5] (version 2.3.3). Konclude is used via its command line interface.

All reasoning performance experiments are run on a machine with an Intel Core i5-3470 Quad-Core processor at 3.2 GHz with 8 GB of RAM. The reasoners were given 5 GB of RAM and the remaining 3 GB were reserved for the operating system (Ubuntu 16.04.04 LTS). The installed Java runtime environment was "OpenJDK Runtime Environment AdoptOpenJDK (build 11.0.4+11)".

Source code used for this work is available online.[6]

5.2 Research Questions

The notion of precompilation raises a number of research questions. Here, we distinguish between two broad categories of such questions. On the one hand, we

[1] https://zenodo.org/record/439510#.XoR4Td-YVhF.
[2] https://www.derivo.de/en/produkte/konclude.html.
[3] http://www.hermit-reasoner.com/.
[4] http://jfact.sourceforge.net/.
[5] https://github.com/stardog-union/pellet.
[6] https://github.com/ckindermann/precompilation.

are interested in the *prevalence* of precompiled entailment sets in practice. On the other hand, we are interested in the *implications* of precompiled entailment sets for practitioners.

To develop a first understanding of precompilation in practice, we focus on entailment sets that are related to entailment sets of standard OWL reasoning services. In particular, we investigate entailment sets revolving around the class hierarchy and class assertions (cf. Sect. 3). We determine to what degree such entailment sets are materialised and to what extent they are redundant and precompiled. Furthermore, we qualify the size of materialised entailment sets relative to an ontology overall size and draw comparisons w.r.t. an ontology's reduction (w.r.t. said entailment set).

Lastly, we shed some light on the practical impact of precompilation by evaluating its effects on reasoning performance. In particular, we investigate reasoning performance with respect to the standard reasoning tasks (i) classification and (ii) realisation.

5.3 Experimental Design

The notion of precompilation is partially predicated on the materialisation of entailment sets and partially on their redundancy. Hence, both the extent of materialisation and redundancy are partial indicators for precompilation. In this work, we investigate precompilation w.r.t. entailment sets revolving around an ontology's class hierarchy (c.f. Sect. 3). Our investigation consists of four distinct experiments that we run over biomedical ontologies as described in Sect. 5.1. The four experiments concern the extent of *materialisation*, *redundancy*, and *precompilation* of entailment sets, as well as the impact of precompilation on *reasoning performance*. In the following, we give a brief description for each of these experiments.

Materialisation. We determine to what extent an ontology consists of atomic axioms and what kinds of atomic axioms are most prevalent. We analyse an ontology's TBox and ABox in the same fashion according to their respective sets defined in Sect. 3. As the entailment sets of transitive reducts and most specific class assertions are of special interest, we shed light on both their materialised and non-materialised proportions. As already mentioned these sets are, in general, not uniquely determined, which makes counting their axioms rather difficult. To avoid over-counting axioms in transitive reducts due to their non-determinism, we adopt the following approach: first, we take an injective function r that returns a representative element $r([\mathsf{A}])$ for each equivalence class $[\mathsf{A}]$. The transitive reduct induced by r is called $r\mathsf{TR}(\mathcal{O})$, and we consider the axiom $r([\mathsf{A}]) \sqsubseteq r([\mathsf{B}]) \in r\mathsf{TR}(\mathcal{O})$ to be materialised in $\mathcal{O}$ if some axiom in $[\mathsf{A}] \sqsubseteq [\mathsf{B}]$ is materialised. Analogously, r induces most specific class assertions $r\mathsf{MSCA}(\mathcal{O})$, and a class assertion $r([\mathsf{A}])(a) \in r\mathsf{MSCA}(\mathcal{O})$ is considered to be materialised if some axiom in $[\mathsf{A}](a)$ is materialised.

For this experiment, we use Konclude to compute an ontology's classification and realisation.[7] When analysing TBoxes, we exclude ontologies with empty TBoxes or empty transitive reducts (TR) of their inferred class hierarchy. Likewise, when analysing ABoxes, we exclude ontologies with empty ABoxes or an empty set of most specific class assertions (MSCA).

Redundancy. We determine to what extent atomic axioms are redundantly contained in ontologies. We test for each axiom α in an ontology $\mathcal{O}$ whether $\mathcal{O} \setminus \{\alpha\} \models \alpha$ holds. As we work with a large number of ontologies that may include many axioms expressed in very expressive DLs, testing atomic axioms individually for redundancy is an expensive operation. Therefore, we configure two timeouts. One timeout, set to two minutes, limits the time a reasoner has to answer an individual redundancy test. A second timeout, set to one hour, limits the time a reasoner has to test all atomic axioms in an ontology. We run this experiment for all three reasoners supported by the OWL API (HermiT, JFact, and Pellet) separately.

Precompilation. We investigate the impact of precompilation on published ontologies by drawing a threefold comparison. We distinguish between the cases of (i) published ontology, (ii) no precompilation, and (iii) minimal precompilation.

For atomic TBox axioms, these three cases are defined as follows: for a given (i) published ontology $\mathcal{O}$ we compute its (ii) reduction $\mathcal{O}^-$ w.r.t. $\mathsf{ICH}(\mathcal{O})$, and (iii) a minimally precompiled ontology defined by $\mathcal{O}^+ = \mathcal{O}^- \cup \mathsf{TR}(\mathcal{O})$ for a transitive reduct that results in a minimal number of added axioms. For atomic ABox axioms, define conditions (i)–(iii) analogously w.r.t. $\mathsf{ICHA}(\mathcal{O})$ and $\mathsf{MSCA}(\mathcal{O})$.

Using Proposition 1, we compare $\mathcal{O}^-$ with $\mathcal{O}$ and $\mathcal{O}^+$ under set difference to measure the impact of precompiling class hierarchy entailment sets on an ontology's size in practice.

Reductions are computed brute-force by iteratively removing redundant axioms. We configure two timeouts as in the redundancy experiment to limit individual reasoning calls and the overall computation. We run this experiment for all three reasoners supported by the OWL API (HermiT, JFact, and Pellet) separately.

Reasoning Performance. We investigate the impact of precompilation on reasoning performance w.r.t. class hierarchy entailment sets. In particular, we time the standard reasoning tasks classification and realisation under three experimental conditions respectively. The three experimental conditions distinguish between the cases of (i) published ontology, (ii) no precompilation, and

[7] Konclude does not compute the realisation of an ontology as it is defined in Sect. 3. Instead, Konclude's `realization` command returns *all* inferable atomic class assertions for an ontology. However, given the inferred class hierarchy of an ontology, one can easily determine the *most specific* class assertions.

(iii) precompilation as defined in the precompilation experiment. We time the classification of $\mathcal{O}$, $\mathcal{O}^-$, and $\mathcal{O}^+$ w.r.t. $\mathsf{ICH}(\mathcal{O})$ and $\mathsf{TR}(\mathcal{O})$ and the realisation w.r.t. $\mathsf{ICHA}(\mathcal{O})$ and $\mathsf{MSCA}(\mathcal{O})$.

In the classification experiment, we remove ABoxes from ontologies to control for confounding effects due to large ABoxes. As there is no analogous operation for the realisation experiment we will discuss confounding factors and limitations in Sect. 7. We run this experiment for all three reasoners supported by the OWL API (HermiT, JFact, and Pellet) separately.

6 Results

6.1 Experiment 1: Materialisation

The experimental conditions as specified in Sect. 5.3 resulted in a total of 394 ontologies (65 atomic, 55 $\mathcal{EL}^{++}$, 274 rich) in the case of TBoxes and 132 (6 atomic, 2 $\mathcal{EL}^{++}$, 124 rich) in the case of ABoxes.

We begin the presentation of results with the materialisation of the transitive reduct (TR)[8] and most specific class assertions (MSCA).

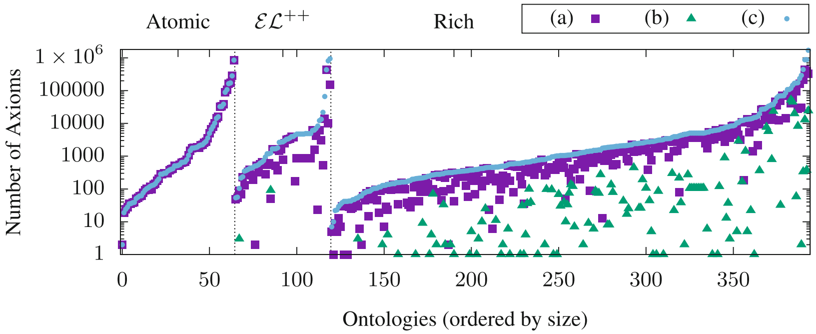
In the case of the TR, we find that 274 (65 atomic, 53 $\mathcal{EL}^{++}$, 156 rich) of the 394 ontologies materialise TR in its entirety. An additional 38 ontologies materialise their TR to at least 99%. Overall, there are only 30 (one $\mathcal{EL}^{++}$, 29 rich) ontologies that materialise their TR to less than 90%. Only 4 of which materialise their TR to less than 50% (the smallest percentage of materialisation is 27%).

In the case of the MSCA, we find that 94 (6 atomic, one $\mathcal{EL}^{++}$, 87 rich) of the 132 ontologies materialise their MSCA in its entirety. An additional 6 ontologies materialise their MSCA to at least 90%. There are only 7 (rich) ontologies that materialise their MSCA to less than 50%; two of which do not materialise any axiom of the MSCA.

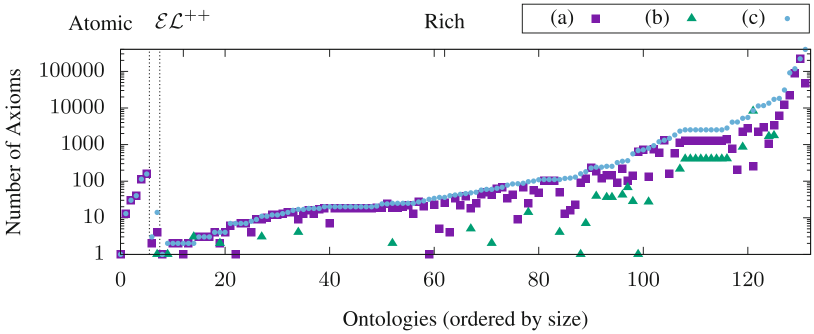
While most ontologies materialise their TR and MSCA to generally high percentages, it is important to relate these percentages to absolute counts. Figure 1A shows absolute counts for the number of axioms in an ontology's TBox, materialised TR axioms, and non-materialised TR axioms. While the total number of non-materialised axioms is below 1000 for most ontologies, there are exceptions. For example, the "The Drug Ontology", shown on index 391, materialises 97% of the TR. Yet, the corresponding number of non-materialised axioms is 12,883.

Figure 1B shows absolute counts for the number of axioms in an ontology's ABox, materialised MSCA axioms, and non-materialised MSCA axioms. We note that ontologies with the huge ABoxes tend to contain a large (absolute) number of axioms from MSCA and materialise their MSCA to 100%; e.g., the RadLex ontology (shown at index 131 in Fig. 1B) has an ABox with 398,016 axioms which includes 46,936 axioms of a materialised MSCA.

[8] Here, we use TR and MSCA as the more abstract concepts that stand for all $\mathsf{TR}(\mathcal{O})$ and $\mathsf{MSCA}(\mathcal{O})$, and remind the reader that we use $r\mathsf{TR}(\mathcal{O})$ and $r\mathsf{MSCA}(\mathcal{O})$ and suitable counting to avoid over-counting these entailments.



(A) Absolute counts for materialised axioms of the TR (a), non-materialised axioms for the TR (b), TBox axioms (c).

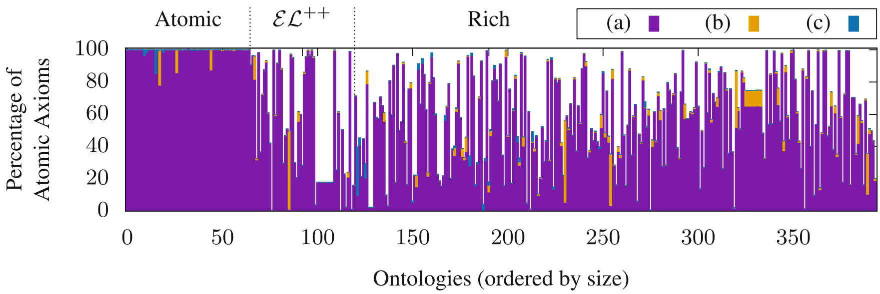


(B) Absolute counts for materialised MSCA axioms (a), non-materialised MSCA axioms (b), and ABox axioms (c).

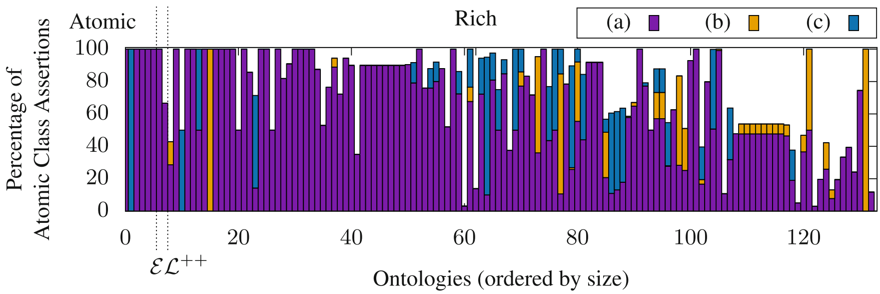
Fig. 1. Size comparison between an ontology's TR and TBox (A), and MSCA and ABox. The legend indicates the drawing order. This order ensures that purple squares (a) cannot hide green triangles (b) nor blue dots (c). Also, a blue dot cannot hide green triangles (b) or purple squares due to its smaller size. (Color figure online)

Figure 1 suggests that the absolute size of materialised TR axioms and MSCA axioms correlate with the size of an ontology's TBox and ABox respectively. However, the logarithmic scale makes it hard to determine visually whether these axioms from the TR and MSCA make up a large proportion of an ontology's TBox or ABox. Figure 2 shows the relative proportions of atomic axioms w.r.t. an ontology's TBox and ABox respectively. It also shows to what extent these axioms are TR axioms or MSCA axioms, short cut axioms, or tautologies.

For TBoxes of non-atomic ontologies, we note that there are both TBoxes with a very small proportion of atomic axioms and TBoxes that consist almost exclusively of atomic axioms. Instances of ontologies with a low proportion of atomic axioms, e.g., at index 112, 118, 212, 275, contain a large number of axioms involving properties (which are non-atomic by our definitions). Interestingly, the proportions of atomic axioms in rich ontologies seem to be almost uniformly distributed in our experimental corpus. We also note, that in most cases, atomic



(A) Percentages are shown for axioms of the TR (a), shot cut axioms (b), and tautologies (c).



(B) Percentages are shown for axioms of the MSCA (a), shot cut axioms (b), and tautologies (c).

Fig. 2. Relative size comparison between an ontology's atomic axioms and TBox (A) and ABox (B).

TBox axioms are indeed TR axioms. However, there are examples where short cuts and tautologies dominate, e.g. at index 121 or 230. While almost all ontologies include a few tautological subclass axioms involving $\top$, the vast majority of axioms from the tautological completion of the TR are not materialised in ontologies.

For ontologies with ABoxes of more than 100 axioms (starting at index 80), we note that the proportion of atomic class assertions tends to decrease. Furthermore, we note that atomic class assertions often contain relatively large numbers of both short cut axioms and tautologies. In particular, we find example ontologies, e.g. at index 77, that materialise all inferred class hierarchy assertions.

6.2 Experiment 2: Redundancy

We report the results of our redundancy experiments for each reasoner by distinguishing for each ontology's ABox and TBox whether (1) redundant atomic axioms could be identified, (2) no redundant atomic axioms could be identified,

(3) the search for redundancies timed out, and (4) whether the case of no redundancies is due to the ABox or TBox being empty. The results are summarised in Table 1.

Table 1. Number of ontologies that contain/do not contain (yes/no) redundant atomic axioms.

Reasoner	ABox			Empty ABox	TBox			Empty TBox	Servicable ontologies
	Yes	No	Timeout		Yes	No	Timeout		
HermiT	59	47	5	241	190	133	24	5	352
JFact	56	46	24	255	200	136	40	5	381
Pellet	41	48	3	227	162	131	21	5	319

While reasoners differ with respect to what ontologies they can service, there is a large overlap of 295 ontologies between all three. Also, while there exist cases in which different reasoners come to different conclusions as to whether a given axiom in a given ontology is redundant, these cases are rare. Therefore, we continue the discussion of experimental results by way of example for HermiT.

We report percentages for the ratio of redundant atomic axioms over all atomic axioms in both TBox and ABox separately. As already mentioned in Sect. 6.1, such percentages may not always give an accurate account of absolute numbers. However, we will defer the discussion of absolute numbers to the experiment on precompilation where ontology reductions are computed.

In TBoxes, the percentage of redundant atomic axioms is rather small for most ontologies. In Fig. 3, on the upward-directed axis, we show to what extent atomic axioms in an ontology's TBox are redundant. For 50 of the 190 ontologies we report that at least 10% of all atomic axioms are redundant. For eight of these ontologies the percentage even surpasses 50%. On the downward-directed axis, we show to what degree redundant axioms are transitive reduct axioms, short cut axioms, or tautologies. We notice that $\mathcal{EL}^{++}$ ontologies contain predominantly redundant short cut axioms. However, in rich ontologies all three kinds of atomic axioms occur as redundant to varying proportions.

In ABoxes, we report comparatively high percentages of redundant atomic axioms. For 31 of the 59 ontologies at least 80% of their atomic class assertions are redundant. Only 20 ontologies contain less than 50% of redundant atomic class assertions. We note, that all three reasoners find redundant atomic class assertions almost exclusively in rich ontologies. In case of Hermit, 57 of the 59 ontologies with redundant atomic class assertions are rich. Lastly, we note that the majority of redundant atomic class assertions are axioms from the MSCA or tautologies. Redundant short cuts occur in 13 ontologies and only make up more than 50% of all redundant atomic axioms in three ontologies.

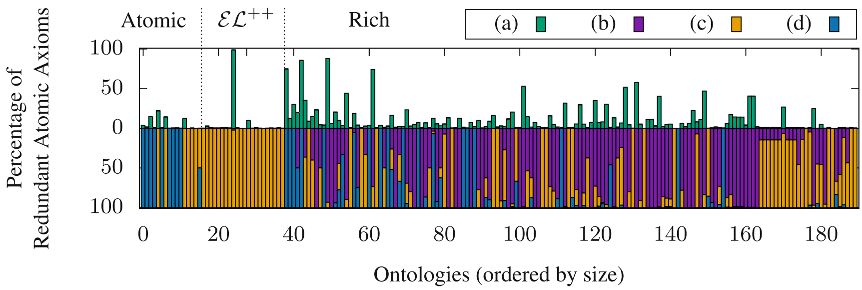


Fig. 3. Prevalence of redundant atomic TBox axioms. Percentages in the upward direction show the ratio of redundant atomic axioms over all atomic axioms of an ontology's TBox (a). Percentages in the downward direction show to what extent redundant atomic axioms are transitive reduct axioms (b), short cut axioms (c), or tautologies (d). Ontologies are ordered by the size of their atomic axioms.

6.3 Experiment 3: Precompilation

We report the results of our precompilation experiment by building on the findings of our materialisation and redundancy experiments. In light of the large proportions of atomic axioms in many ontologies of reasonably large size, removing or adding even small percentages of redundant entailment sets may have a significant impact in practice. For example, in the "FoodOn" ontology, we identified a precompiled set of atomic axioms in its TBox that makes up only 1% of all its atomic axioms. Removing this set from the ontology changed the ontology's overall size only by 0.6%. Yet, more than 100 axioms have been removed.

The results of our precompilation experiment show that such cases are not uncommon. Figure 4 shows absolute counts for the number of axioms in an ontology's TBox, the number of removed axioms in comparison with its reduction (as computed by HermiT), the minimal number of axioms required to add to the reduction so that the TR is materialised in its entirety. We note that 34 of the 190 ontologies (c.f. Fig. 3 in Sect. 6.2) contain precompiled sets of more than 100 atomic axioms.[9] We also note, that there are quite a number of ontologies for which $\mathcal{O}^+ \setminus \mathcal{O}^-$ is rather large. For example, in case of the "Non-coding RNA Ontology", adding a minimal number of axioms to its reduction amounts to the addition of 7659 axioms.

Lastly, we mention that the results of the precompilation experiment for ABoxes are similar to the results we report for TBoxes.

[9] We remind the reader of Proposition 1 according to which a reduction of an ontology w.r.t. an entailment set uniquely identifies a precompiled set of axioms in that ontology.

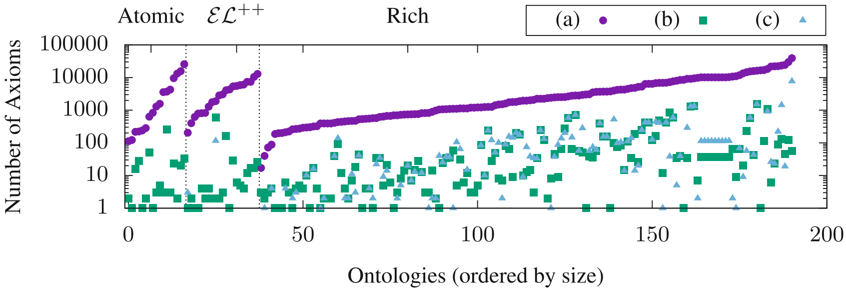


Fig. 4. Size comparison between $\mathcal{O}, \mathcal{O}^-$, and $\mathcal{O}^+$ w.r.t. TBox changes. Absolute counts for axioms of an ontology $\mathcal{O}$'s TBox (a), axioms of $\mathcal{O}$'s precompiled set associated with $\mathcal{O}^-$ (b), the minimal number of axioms to add to $\mathcal{O}^-$ so as to fully materialise its TR (c).

6.4 Experiment 4: Reasoning Performance

We report the results of our reasoning performance in relation to ontologies for which a reasoner spent more than 10 s to solve a reasoning task. For HermiT there are 12 such ontologies for classification and 21 for realisation. For JFact we have 11 and 2 respectively, and for Pellet we have 3 and 4. All results were found for rich ontologies.

Our experiments indicate that precompilation can affect classification times both positively as well as negatively depending on which reasoner is used on what ontology.

HermiT is either consistently unaffected by precompilation of an ontology's TR or seems to benefit. Table 2 summarises the classification times for four of five cases in which the classification times noticeably improve as a result of precompilation. The table also includes the one exception, namely the Immunogenetics Ontology (imgt), for which HermiT's performance suffers under the effect of precompilation.

JFact on the other hand is consistently affected negatively by precompilation. For 8 of the 11 ontologies, JFact's classification time increases considerably. In one case the time increases from 3 min to more than 15 – in another case, the time increases from 16 min to more than 40. There is only one ontology that did not incur performance degradation with JFact under precompilation.

Pellet shows improved reasoning times for one ontology, no effect for another, and yet slightly more volatile behaviour in terms of minimal and maximal classification times that produce similar averages as the no-precompilation condition.

Reasoning behaviour with respect to realisation appears to be largely unaffected by precompilation of the MSCA. JFact was negatively affected by precompilation for one ontology resulting in an increased time from 10 s (no-precompilation) to 30 s (precompilation). Pellet's reasoning time, on the other hand, improved for one ontology from 20 s (no-precompilation) down to 4 s (precompilation).

Table 2. HermiT Classification time for five experimental runs

Ontology	bt-biotop			fb-cv			imgt			ntdo			stato		
Condition	Min.	Avg.	Max.	Min.	Avg.	Max.	Min.	Avg.	Max.	Min.	Avg.	Max.	Min.	Avg.	Max.
$\mathcal{O}^-$	146 s	171 s	206 s	38 s	46 s	52 s	19 s	19 s	20 s	190 s	259 s	322 s	415 s	438 s	456 s
$\mathcal{O}$	96 s	116 s	152 s	11 s	13 s	16 s	77 s	79 s	82 s	142 s	174 s	224 s	276 s	290 s	305 s
$\mathcal{O}^+$	84 s	102 s	131 s	12 s	13 s	15 s	72 s	77 s	80 s	137 s	144 s	155 s	283 s	300 s	306 s

7 Discussion

We find that a good portion of ontologies indexed in BioPortal contains their TR in its entirety. However, consistent with the hypothesis of low redundancy formulated in [6], we find that most materialised axioms from the ICH are indeed not redundant and hence not precompiled. Yet, in case of large ontologies, it is important to keep in mind that even small percentages of redundant axioms may in fact correspond to a large absolute number of axioms. Concerns about such redundancies have been raised on the grounds of their informational value for ontology curators as well as tools [10,13].

Furthermore, the observed differences in reasoning times by a factor of three or more may be of practical relevance. Even in case of only a few seconds, such differences may become noticeable if ontologies need to be classified frequently or in bulk.

Limitations. We have limited the scope of our investigation to the biomedical domain. This design choice is primarily motivated by the view that precompilation of entailment sets is a "fundamental step in the release process for biomedical ontologies [7]." While the notion of precompilation is independent of a particular domain, we are unaware of strong beliefs about precompilation in other domains. Apart from this, BioPortal is a large corpus of actively maintained ontologies that are highly heterogeneous in terms of size and complexity [5,9]. Thus, the generally small effect size of precompilation w.r.t. BioPortal is unlikely to change w.r.t. other corpora of comparable size and complexity.

By using a one-hour timeout in combination with a straightforward method to identify redundancies and computing ontology reductions, we may have systematically excluded computationally challenging ontologies from our reasoning evaluation. Note, however, that our chosen approach proves to be sufficient for the majority of ontologies in our corpus. For example, in the case of HermiT, we only exclude 24 out of 352, i.e., less than 7% of serviceable ontologies. Also note, that a successful treatment of these excluded ontologies would not change the quintessence of our two primary observations. Namely, that precompilation has no effect for the vast majority of ontologies and that precompilation can have both positive as well as negative effects in specific instances.

Lastly, we need to point out that our reasoning experiments were primarily designed to investigate the *impact* of precompilation *in practice*. We did not investigate the *potential* of precompilation to affect reasoning performance *in* *general*. Although we controlled for long classification times due to large ABoxes, we did not control for factors that could (in theory) interact with precompilation. For example, it is conceivable that realisation times depend on both precompilation of the MSCA as well as precompilation of the TR. Similarly, one can speculate that different kinds of precompiled entailment sets would affect each others impact w.r.t. reasoning performance.

Future Work. Given the observed effects of precompilation on reasoning performance in a few instances, an understanding of cause and effect would be valuable. For this purpose, we plan a study on computationally challenging ontologies (not restricted to a domain). Here, we give a brief description of crucial points for future experimentation.

To investigate the potential impact of precompilation, ontology reductions w.r.t. to the whole ontology need to be considered. Since the number of such reductions is (in theory) exponential, stochastic sampling for exploring this search space may be sensible. Similarly, different precompilations of an entailment set, e.g., via minimal but non-unique representations, need to be analysed (also possibly by stochastic sampling).

A straightforward approach to compute ontology reductions is unlikely to be sufficient for a study on computationally challenging ontologies. Thus, optimisations and approximation techniques are needed; especially for large scale experiments.

A detailed analysis of used algorithms and concrete implementations of reasoners is necessary to develop an understanding of reasoner specific behaviour. Software profiling techniques may provide useful information for pinpointing implementation-specific factors contributing to effects of precompilation on reasoning time.

Ultimately, the potential impact of precompilation on reasoning performance involves three independent factors: an ontology, an entailment set, and a reasoner. Thus, a full investigation of this impact will need to examine all three factors as well as their potential interactions.

Related Work. Precompilation and materialisation of entailment sets is discussed in a range of settings. First, related but different precompilation approaches involve rewriting an ontology into a certain normal form to make subsequent tasks, including reasoning, easier [2,4]. Here, we focus on *extending* an ontology with entailed axioms.

Second, materialisation is used to compensate for shortcomings of tools. In [12], the materialisation of entailment sets is proposed to mitigate limitations of incomplete reasoners. The main idea is to determine entailment sets $\mathcal{R}$ that function as a *repair* without which an incomplete reasoner would fail to derive answers to some queries.

Third, materialisation is used to preserve an ontology's entailments when it is translated into a less expressive description logic in [11]. We could say that materialisation, in this setting, compensates for lack of expressive power which, in turn, can be motivated by reasoner or tool performance requirements or other reasons like readability.

Fourth, materialisation of TBox entailments has been used to improve the performance of (ABox) query answering in a range of settings, e.g., in [8].

Finally, while there are numerous surveys of properties of existing ontologies, there are none—to the best of our knowledge—that investigate the extent of materialisation.

8 Conclusion

We find that biomedical ontologies materialise class hierarchy entailment sets in both their TBox and Abox. While these entailment sets are to large proportions a substantive part of the ontology and cannot be removed without changing the ontology's meaning, there exist redundant subsets in ontologies that are of non-trivial size. Likewise, adding the TR to an ontology may result in a non-trivial increase in size. While our experiments on reasoning performance suggest that precompilation is inconsequential in most cases, there are instances where this practice can have a noticeable impact, both positive as well as negative, that depends on the used reasoner for a given ontology.

Overall, we conclude that the practice of precompilation has to be treated with due diligence. Adding entailment sets in an automated manner can have a significant impact on both an ontology's size and its usability in practice. Whether the precompilation of an entailment set provides its desired beneficial effects needs to be tested on a case by case basis. Tool support to facilitate such testing will be of great value moving forward.

References

1. Bail, S., Parsia, B., Sattler, U.: Extracting finite sets of entailments from OWL ontologies. In: Description Logics. CEUR Workshop Proceedings, vol. 745. CEUR-WS.org (2011)
2. Bienvenu, M.: Prime implicate normal form for ALC concepts. In: AAAI, pp. 412–417. AAAI Press (2008)
3. Cuenca Grau, B., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P.F., Sattler, U.: OWL 2: the next step for OWL. J. Web Semant. **6**(4), 309–322 (2008)
4. Furbach, U., Günther, H., Obermaier, C.: A knowledge compilation technique for ALC Tboxes. In: FLAIRS Conference. AAAI Press (2009)
5. Gonçalves, R.S., Matentzoglu, N., Parsia, B., Sattler, U.: The empirical robustness of description logic classification. In: Description Logics. CEUR Workshop Proceedings, vol. 1014, pp. 197–208. CEUR-WS.org (2013)
6. Grimm, S., Wissmann, J.: Elimination of redundancy in ontologies. In: Antoniou, G., et al. (eds.) ESWC 2011. LNCS, vol. 6643, pp. 260–274. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21034-1_18
7. Jackson, R.C., Balhoff, J.P., Douglass, E., Harris, N.L., Mungall, C.J., Overton, J.A.: ROBOT: a tool for automating ontology workflows. BMC Bioinform. **20**(1), 407:1–407:10 (2019)

8. Liu, Y., McBrien, P.: SPOWL: spark-based OWL 2 reasoning materialisation. In: Proceedings of the 4th ACM SIGMOD Workshop on Algorithms and Systems for MapReduce and Beyond, pp. 3:1–3:10. ACM (2017)
9. Matentzoglu, N., Bail, S., Parsia, B.: A snapshot of the OWL web. In: Alani, H., et al. (eds.) ISWC 2013. LNCS, vol. 8218, pp. 331–346. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41335-3_21
10. Mougin, F.: Identifying redundant and missing relations in the gene ontology. In: MIE. Studies in Health Technology and Informatics, vol. 210, pp. 195–199. IOS Press (2015)
11. Pan, J.Z., Thomas, E.: Approximating OWL-DL ontologies. In: AAAI, pp. 1434–1439. AAAI Press (2007)
12. Stoilos, G., Grau, B.C.: Repairing incomplete reasoners. In: Description Logics. CEUR Workshop Proceedings, vol. 745. CEUR-WS.org (2011)
13. Xing, G., Zhang, G., Cui, L.: FEDRR: fast, exhaustive detection of redundant hierarchical relations for quality improvement of large biomedical ontologies. BioData Min. **9**, 31 (2016)