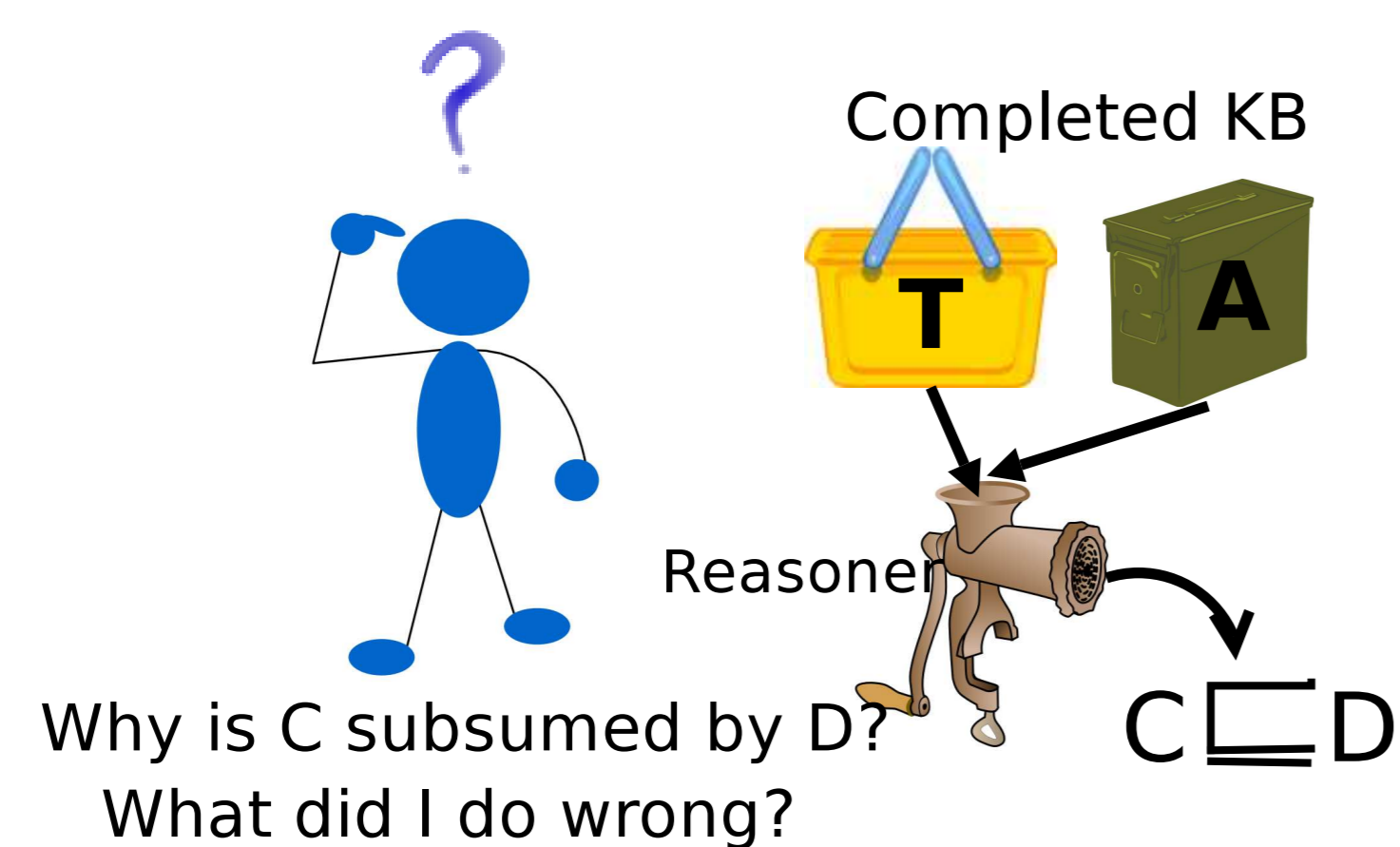




1. Abstract

Knowledge base completion is a method for extending both the terminological and assertional part of a Description Logic knowledge base by using information provided by a domain expert. It ensures that the extended knowledge base is complete w.r.t. a fixed interpretation in a certain, well-defined sense. Here we consider the problem of explaining user errors in knowledge base completion. We show that for this setting, the problem of deciding the existence of an explanation within a specified cardinality bound is NP-complete, and the problem of counting explanations that are minimal w.r.t. set inclusion is #P-complete. We also provide an algorithm that computes one minimal explanation by performing at most polynomially many subsumption tests.

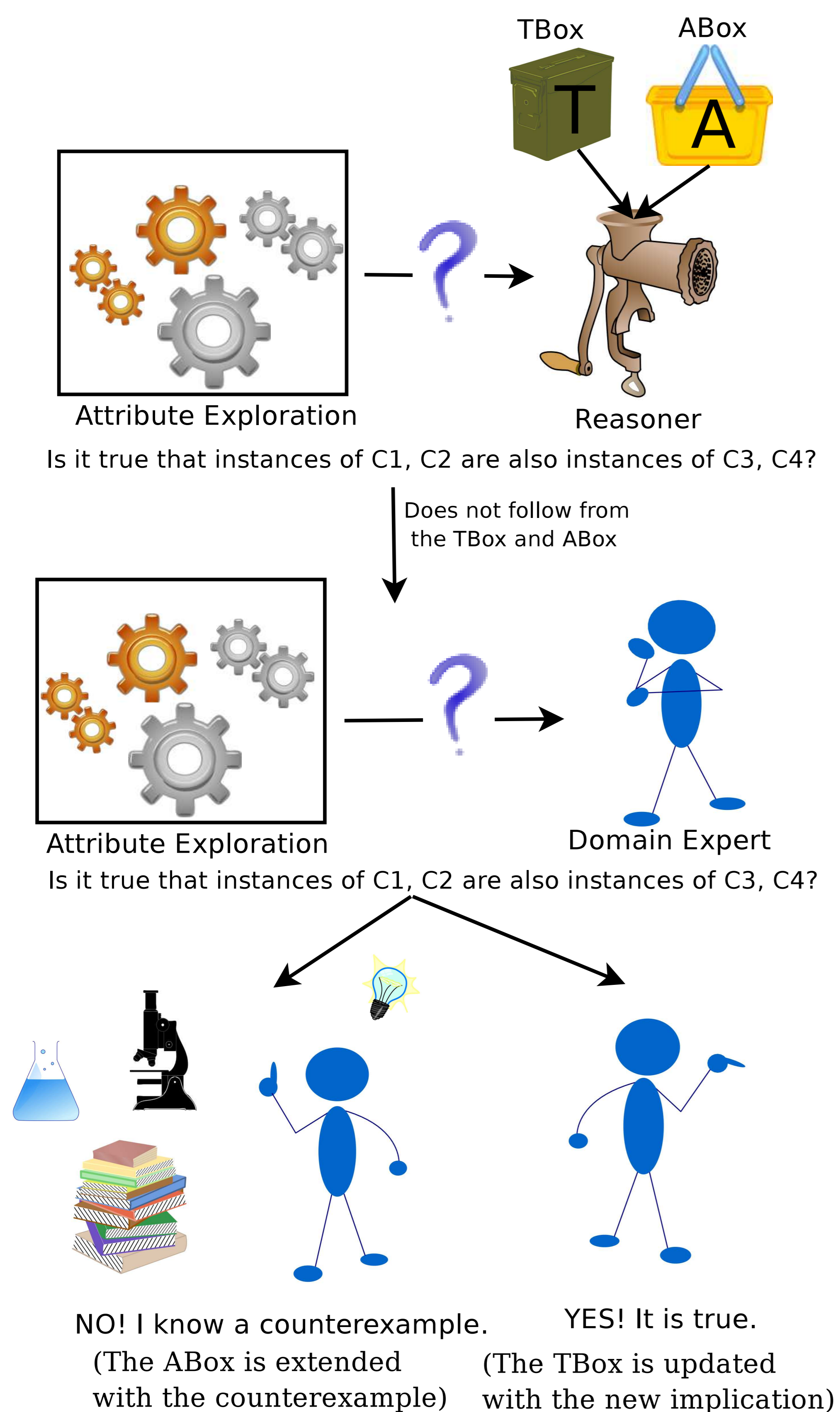
2. Introduction



In [1, 2] we have developed a method for supporting the ontology engineer in checking whether an ontology contains all the relevant information about the application domain, and extending the ontology appropriately if this is not the case. The method achieves this by asking the ontology engineer questions of the form: “is it true that instances of the concepts C_1, \dots, C_n are also instances of the concepts D_1, \dots, D_m ?” When such a question is asked, the expert is expected to either confirm or reject it. Once all such questions are answered, the knowledge base will be complete in a certain sense.

Our first experiments with a prototype implementation of the method showed that during completion, the ontology engineer sometimes by mistake confirms a wrong question, in which case an axiom that actually does not hold in the application domain is added to the TBox. If she does not notice it immediately, as a result the completed knowledge base will have unwanted consequences. In the present work we consider the problem of pinpointing the axioms added during completion, which lead to certain unwanted consequences.

3. Knowledge Base Completion



4. Explaining user errors

Lemma 0.1 Let \mathcal{L} be a Duquenne Guigues Base. Every implication $L \rightarrow R$ in \mathcal{L} has the following properties:

- L is closed under the implications $\mathcal{L} \setminus \{L \rightarrow R\}$, i.e., $L = \mathcal{L}'(L)$ where \mathcal{L}' is $\mathcal{L} \setminus \{L \rightarrow R\}$ and $\mathcal{L}(\cdot)$ denotes implicational closure,
- $L \cup R$ is closed under the implications $\mathcal{L} \setminus \{L \rightarrow R\}$, i.e., $L \cup R = \mathcal{L}'(L \cup R)$.

The following example shows that a subsumption relation can have exponentially many minimal explanations in a TBox that obeys the property in Lemma 0.1.

Example 0.2 Consider the TBox

$$\mathcal{T} := \{X \sqcap B_{i-1} \sqsubseteq P_i \sqcap Q_i, Y \sqcap P_i \sqsubseteq B_i, Y \sqcap Q_i \sqsubseteq B_i \mid 1 \leq i \leq n\}.$$

Assume that all axioms in \mathcal{T} are refutable. It is not difficult to see that none of the left hand sides is contained in another left hand side or in the union of left and right hand sides of another axiom, i.e., it obeys the property mentioned in Lemma 0.1. Moreover its size is linear in n , and it has 2^n minimal subsets that explain the axiom $B_0 \sqcap X \sqcap Y \sqsubseteq B_n$ since for each i , $1 \leq i \leq n$, B_i can be generated by the axiom $Y \sqcap P_i \sqsubseteq B_i$ or by $Y \sqcap Q_i \sqsubseteq B_i$.

Problem: MINIMUM CARDINALITY EXPLANATION

Input: A Horn TBox \mathcal{T} satisfying the properties in Lemma 0.1, sets L and R of concept names occurring in \mathcal{T} such that $\sqcap L \sqsubseteq_{\mathcal{T}} \sqcap R$, a natural number n .

Question: Is there an explanation of $\sqcap L \sqsubseteq \sqcap R$ in \mathcal{T} with cardinality less than or equal to n , i.e., is there a set of axioms $\mathcal{T}' \subseteq \mathcal{T}$ such that $\sqcap L \sqsubseteq_{\mathcal{T}'} \sqcap R$ and $|\mathcal{T}'| \leq n$?

Theorem 0.3 MINIMUM CARDINALITY EXPLANATION is NP-complete.

Proof. Hardness by reduction from HITTING SET.

Problem: #MINIMAL EXPLANATION

Input: A Horn TBox \mathcal{T} satisfying the properties in Lemma 0.1, sets L and R of concept names occurring in \mathcal{T} such that $\sqcap L \sqsubseteq_{\mathcal{T}} \sqcap R$.

Output: Number of all minimal explanations of $\sqcap L \sqsubseteq \sqcap R$ in \mathcal{T} , i.e., $|\{\mathcal{T}' \subseteq \mathcal{T} \mid \sqcap L \sqsubseteq_{\mathcal{T}'} \sqcap R \text{ and } \forall \mathcal{T}'' \subsetneq \mathcal{T}'. \sqcap L \not\subseteq_{\mathcal{T}''} \sqcap R\}|$.

Theorem 0.4 #MINIMAL EXPLANATION is #P-complete.

Proof. Hardness by reduction from #HITTING SET.

Despite these negative results, it is not difficult to find one minimal explanation with at most polynomially many subsumption tests. We can just start with the whole set of axioms obtained from the completion process, iterate over these axioms and eliminate an axiom if the remaining ones still have the consequence in question. It is formally described in Algorithm ??.

1. **Input:** The set of axioms \mathcal{T} obtained from completion, and sets of concept names L and R s.t. $\sqcap L \sqsubseteq_{\mathcal{T}} \sqcap R$.
2. $\mathcal{T}' := \mathcal{T}$
3. **for all** $t \in \mathcal{T}'$ **do**
4. **if** $\sqcap L \sqsubseteq_{\mathcal{T}' \setminus \{t\}} \sqcap R$ **then** $\{\text{if } \mathcal{T}' \setminus \{t\} \text{ is still an explanation}\}$
5. $\mathcal{T}' := \mathcal{T}' \setminus \{t\}$
6. **end if**
7. **end for** \mathcal{T}'

9. Conclusion

As future work we are going to investigate the complexity of the problem of computing all minimal explanations in the KB completion setting. If the following decision problem is not in P, there cannot be an output-polynomial algorithm that performs this task: given a set of minimal explanations, decide whether there exists another minimal explanation that does not occur in this set. In [3] it was shown that in the setting where the set of axioms is arbitrary Horn clauses, this problem is NP-complete.

References

- [1] F. Baader, B. Ganter, U. Sattler, and B. Sertkaya. Completing description logic knowledge bases using formal concept analysis. LTCS-Report LTCS-06-02, Chair for Automata Theory, Institute for Theoretical Computer Science, Dresden University of Technology, Germany, 2006. See <http://lat.inf.tu-dresden.de/research/reports.html>.
- [2] F. Baader, B. Ganter, B. Sertkaya, and U. Sattler. Completing description logic knowledge bases using formal concept analysis. In M. M. Veloso, editor, *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI'07)*, pages 230–235. AAAI Press, 2007.
- [3] F. Baader, R. Peñaloza, and B. Suntisrivaraporn. Pinpointing in the description logic \mathcal{EL}^+ . In J. Hertzberg, M. Beetz, and R. Englert, editors, *Proceedings of the 30th German Conference on Artificial Intelligence (KI2007)*, volume 4667 of *Lecture Notes in Artificial Intelligence*, pages 52–67. Springer-Verlag, 2007.