Optimization and Empirical Evaluation of the CATS ABox Abduction Solver (Extended Abstract)

Janka Boborová[1], Jakub Kloc[2], Martin Homola[1] and Júlia Pukancová[1]

[1]Comenius University in Bratislava, Mlynská dolina, 842 41 Bratislava, Slovakia
[2]SOFTEC, Einsteinova 33, 851 01 Bratislava, Slovakia

Abstract

CATS is a JFact-based ABox abduction solver that implements three classical complete approaches: Reiter's minimal hitting set algorithm and its more recent variants HS-Tree and RC-Tree. In addition, it provides fast but incomplete methods, namely QuickXplain and MergeXplain, as well as hybrid variants in which the complete algorithms are "boosted" using MergeXplain. This hybrid strategy is theoretically motivated and is expected to provide significant performance advantages for specific classes of inputs, as suggested by earlier preliminary results. In this work, we report on the recently implemented optimizations and on the first empirical evaluation conducted on three real-world ontologies that enables to compare all eight implemented algorithms. Among other insights, it shows a decisive advantage of the hybrid methods over all three baselines on all classes of inputs.

Keywords

ABox abduction, abductive reasoning, description logics, evaluation

1. Introduction

ABox abduction [1] addresses the task of explaining observations that are not entailed by a given ontology by identifying missing or hypothetical ABox assertions whose addition would render the observation derivable. Formally, an ABox abduction problem is given as a tuple $(\mathcal{K}, \mathcal{O}, \mathrm{Abd})$, where $\mathcal{K}$ is a description logic (DL) knowledge base, $\mathcal{O}$ is a finite set of observed ABox assertions, and Abd is a designated set of abducible assertions. An explanation is a subset $\mathcal{E} \subseteq \mathrm{Abd}$ such that $\mathcal{K} \cup \mathcal{E} \models \mathcal{O}$, typically subject to additional requirements, including consistency of $\mathcal{K} \cup \mathcal{E}$, relevance ($\mathcal{E} \not\models \mathcal{O}$), and minimality with respect to set inclusion. It is well established that ABox abduction can be reduced to the computation of minimal hitting sets of collections of conflict sets derived from models of $\mathcal{K} \cup \neg\mathcal{O}$[1], yielding a logic-agnostic formulation that relies solely on black-box reasoning (provided the access to the relevant part of the model).

From a practical perspective, ABox abduction plays a central role in a wide range of applications, including ontology debugging [3] and repair [4], medical diagnosis [5, 6, 7], system and fault diagnosis [8], and explanation generation for knowledge-based and machine-learning systems [9]. Such applications require methods that are both complete – capable of enumerating all minimal explanations – and sufficiently efficient to operate on real-world ontologies, thus motivating the development of scalable abductive reasoning techniques.

2. CATS ABox Abduction Solver

CATS [10] is a fully modular Java-based solver for DL ABox abduction. It uniformly integrates multiple abductive reasoning algorithms under a shared interface and delegates consistency checking and model extraction to an external DL reasoner (JFact). This design ensures independence from the concrete DL

DL 2026: 39th International Workshop on Description Logics, July 17–19, 2026, Lisbon, Portugal
boborova@fmph.uniba.sk (J. Boborová); kloc4@uniba.sk (J. Kloc); homola@fmph.uniba.sk (M. Homola); pukancova@fmph.uniba.sk (J. Pukancová)
https://dai.fmph.uniba.sk/~homola/ (M. Homola); https://dai.fmph.uniba.sk/~pukancova/ (J. Pukancová)
0009-0002-7487-9962 (J. Boborová); 0000-0001-6384-9771 (M. Homola); 0009-0001-3505-0716 (J. Pukancová)

[1]If $\mathcal{O} = \{O\}$ is a singleton, then $\neg\mathcal{O} = \{\neg O\}$, otherwise we rely on the known reduction to single observation [2].

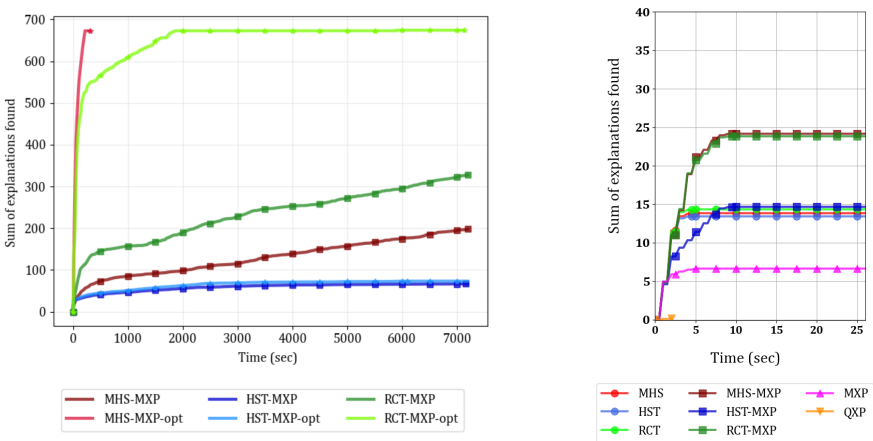
and enables support for expressive ontologies up to OWL 2/$\mathcal{SROIQ}$, while also allowing direct and fair empirical comparison of different algorithms.

Three complete baseline algorithms based on hitting-set computation are implemented: Minimal Hitting Set (MHS) [11], Hitting Set Tree (HST) [12], and RC-Tree (RCT) [13]. All three enumerate minimal explanations by constructing variants of an HS-tree whose nodes represent partial hitting sets labelled by conflict sets obtained from models. MHS follows a classical breadth-first construction. HST improves enumeration efficiency by generating all candidate subsets while ensuring that each combination of abducibles is constructed at most once. RCT further reduces redundancy by dynamically preventing the creation of equivalent branches using exclusion sets. Despite these improvements, all baseline algorithms remain limited by the intrinsic computational complexity of complete minimal hitting-set enumeration.

In addition, two incomplete divide-and-conquer algorithms are implemented: QuickXplain (QXP) [14] and MergeXplain (MXP) [15]. These methods are designed to rapidly identify one or more minimal explanations but do not guarantee completeness and may fail to enumerate all solutions.

The primary contribution of the solver lies in its hybrid algorithms – MHS-MXP [16], HST-MXP, and RCT-MXP – which combine complete hitting-set search with the look-ahead capabilities of MXP. The hybrids replace classical consistency checks in the search tree with MXP calls that can directly extend partial candidate explanations and identify branches that cannot lead to valid solutions. This design is motivated by the observation that baseline algorithms often expend significant effort exploring unproductive regions of the search space, while MXP can provide early global insight. All hybrid variants are shown to retain soundness and completeness while substantially accelerating explanation discovery in practice.

3. Optimizations



(a) (b)

Figure 1: (a) Impact of optimizations; (b) Evaluation results. At each time point, the curves show the average number of desired explanations found up to that time across all runs included in the corresponding evaluation.

A number of optimizations were incorporated[2] to improve the practical performance of both baseline and hybrid algorithms, with particular emphasis on reducing the cost of reasoning calls and limiting unnecessary exploration of the search space. The previous version of CATS [10] already incorporated several implementation-level optimizations, most notably *model caching* and *memory-efficient HS-tree management*. Extracted models are cached and reused for node labelling, reducing the number of expensive reasoning calls. HS-tree-based algorithms, specifically MHS(-MXP) and HST(-MXP), employ

[2]The CATS source code is publicly available at https://github.com/Comenius-Abduction-Team/CATS-Abduction-Solver. All optimizations described in this section are included in release 2.1.2.

memory-efficient tree management by storing only unprocessed nodes instead of the complete search tree, significantly reducing memory consumption.

In this work, we investigate several additional optimization strategies aimed at further improving the efficiency of explanation computation: *(i)* The number of consistency checks was reduced to a minimum by identifying points where they are unnecessary. *(ii)* Models are preferentially selected based on the size of the induced conflict set. Smaller conflict sets yield lower branching factors, resulting in more compact HS-trees and faster convergence. *(iii)* An optimization that specifically targets a known weakness [16] of MXP-based hybrid approaches in conflict-dense abducible spaces: Instead of invoking MXP only once per node, the solver applies it iteratively to different partitions of the remaining abducibles: first to positive assertions, then to negative assertions, and finally to the full set. This staged strategy often enables earlier identification of explanations and improves robustness in unfavourable search spaces.

We evaluated the impact of the newly introduced optimizations on individual algorithms using selected inputs of varying difficulty constructed over the LUBM ontology [10]. Based on these experiments, we determined for each algorithm which optimizations are effective and compatible with it. The algorithms are then automatically executed in CATS with the optimizations that best suit them.

Optimization *(i)* proved crucial and significantly improved the efficiency of all HS-tree-based algorithms. Optimization *(ii)* was not effective for HST(-MXP) and occasionally even led to performance degradation. Although it was effective for RCT(-MXP), this algorithm uses model ordering, and the optimization interferes with it, potentially causing premature termination before all explanations are found. Until this issue is resolved, it is not suitable to use this optimization with RCT(-MXP); at present, it is effective and compatible only with MHS(-MXP). Optimization *(iii)* was shown to be beneficial for all hybrid algorithms.

Figure 1 (a) compares the performance of hybrid algorithms before and after enabling their respective compatible optimizations on selected LUBM inputs of varying difficulty.

4. Evaluation

A principled **methodology**[3] was developed to construct realistic benchmarks for DL ABox abduction. Three real-world ontologies were selected from the ORE 2015 Reasoner Competition corpus, chosen to represent different sizes and expressivity levels (two $\mathcal{EL}$++ ontologies and one $\mathcal{ALCHON}$ ontology).

Observations were derived from actual logical consequences of the ontologies, restricted to atomic and negated atomic ABox assertions. To introduce abductive problems with guaranteed solutions, the ontologies were modified by removing ABox justifications responsible for deriving the selected consequences.

Using this methodology, diverse classes of inputs were generated, distinguished by the minimum explanation size (1–5), the number of minimal explanations (ranging from less than 10 to more than 100 per input). Each input was paired with six sets of abducibles with different size (50, 100 and 200), each size in two variants (random, and conflicting, the latter comprising of pairs of complementary literals to emulate the conflict-dense abducible space mentioned above).

Figure 1 (b) summarizes the overall performance across all generated inputs. We observe a clear advantage of hybrid algorithms over baseline approaches. On average, MHS-MXP and RCT-MXP compute significantly more explanations and discover them earlier than any baseline algorithm, with MHS-MXP achieving the best overall performance. In contrast, HST-MXP consistently underperforms due to its less compact search structure and frequent memory exhaustion.

Among baseline algorithms, MHS, HST, and RCT exhibit broadly similar behaviour, with RCT achieving a modest but consistent advantage. As expected, the incomplete algorithms MXP and QXP return fewer explanations and mainly serve as reference points.

The advantage of hybrid algorithms increases with explanation length and abducible set size. Although conflict-dense abducible spaces reduce the relative margin, MHS-MXP and RCT-MXP remain

[3]A preliminary version of the methodology was reported at the previous DL Workshop [17].

dominant on average, representing a substantial improvement over earlier evaluations in unfavourable configurations.

5. Conclusions

The results confirm that hybrid abduction algorithms effectively reconcile completeness with practical efficiency in DL ABox abduction. By integrating MXP-based look-ahead into complete hitting-set search, MHS-MXP and RCT-MXP substantially outperform classical baseline methods, particularly for large abducible sets and long explanations. The modular solver and systematically constructed benchmarks enable a comprehensive empirical comparison of general-purpose DL ABox abduction techniques. From a practical perspective, MHS-MXP emerges as the default choice for scalable complete abduction, with RCT-MXP providing a robust alternative when redundancy control is critical.

Acknowledgments

This research was sponsored by Slovak Republic under the SRDA grant APVV-23-0292 (DyMAX) and VEGA grant no. 1/0630/25 (eXSec). J. Boborová was supported by a DL student grant and Grant No. UK/1353/2026 from Comenius University Bratislava.

Declaration on Generative AI

During the preparation of this work, the authors used Microsoft Copilot (powered by GPT-5) for: Paraphrasing and rewording, Abstract drafting, Improving writing style, Grammar and spelling checking, and Formatting assistance. After using these services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] C. Elsenbroich, O. Kutz, U. Sattler, A case for abductive reasoning over ontologies, in: Proceedings of the OWLED*06 Workshop on OWL: Experiences and Directions, Athens, GA, US, volume 216 of *CEUR-WS*, CEUR-WS.org, 2006. URL: https://ceur-ws.org/Vol-216/submission_25.pdf.

[2] J. Pukancová, M. Homola, ABox abduction for description logics: The case of multiple observations, in: Proceedings of the 31st International Workshop on Description Logics, Tempe, Arizona, US, volume 2211 of *CEUR-WS*, CEUR-WS.org, 2018. URL: https://ceur-ws.org/Vol-2211/paper-31.pdf.

[3] K. Schekotihin, P. Rodler, W. Schmid, OntoDebug: Interactive ontology debugging plug-in for Protégé, in: Foundations of Information and Knowledge Systems - 10th International Symposium, FoIKS 2018, Budapest, Hungary, May 14-18, 2018, Proceedings, volume 10833 of *LNCS*, Springer, 2018, pp. 340–359. doi:10.1007/978-3-319-90050-6_19.

[4] F. Wei-Kleiner, Z. Dragisic, P. Lambrix, Abduction framework for repairing incomplete $\mathcal{EL}$ ontologies: Complexity results and algorithms, in: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada., AAAI Press, 2014, pp. 1120–1127. doi:10.1609/AAAI.V28I1.8858.

[5] S. M. Rashid, J. P. McCusker, D. M. Gruen, O. Seneviratne, D. L. McGuinness, A concise ontology to support research on complex, multimodal clinical reasoning, in: The Semantic Web - 20th International Conference, ESWC 2023, Hersonissos, Crete, Greece, May 28 - June 1, 2023, Proceedings, volume 13870 of *Lecture Notes in Computer Science*, Springer, 2023, pp. 390–407. doi:10.1007/978-3-031-33455-9_23.

[6] C. Martini, Abductive reasoning in clinical diagnostics, in: L. Magnani (Ed.), Handbook of Abductive Cognition, Springer, 2023, pp. 467–479. doi:10.1007/978-3-031-10135-9_13.

[7] M. Obeid, Z. Obeid, A. Moubaiddin, N. Obeid, Using description logic and ABox abduction to capture medical diagnosis, in: Advances and Trends in Artificial Intelligence. From Theory to Practice - 32nd International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2019, Graz, Austria, July 9-11, 2019, Proceedings, volume 11606 of *LNCS*, Springer, 2019, pp. 376–388. doi:10.1007/978-3-030-22999-3_33.

[8] T. Hubauer, C. Legat, C. Seitz, Empowering adaptive manufacturing with interactive diagnostics: A multi-agent approach, in: Advances on Practical Applications of Agents and Multiagent Systems – 9th International Conference on Practical Applications of Agents and Multiagent Systems, PAAMS 2011, Salamanca, Spain, volume 88 of *Advances in Intelligent and Soft Computing*, Springer, 2011, pp. 47–56. doi:10.1007/978-3-642-19875-5_6.

[9] A. Ignatiev, N. Narodytska, J. Marques-Silva, On relating explanations and adversarial examples, in: H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada, 2019, pp. 15857–15867. URL: https://proceedings.neurips.cc/paper/2019/hash/7392ea4ca76ad2fb4c9c3b6a5c6e31e3-Abstract.html.

[10] J. Kloc, J. Boborová, M. Homola, J. Pukancová, CATS solver: The rise of hybrid abduction algorithms, in: L. Tendera, Y. Ibáñez-García, P. Koopmann (Eds.), Proceedings of the 38th International Workshop on Description Logics - DL 2025, Opole, Poland, September 3-6, 2025, volume 4091 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2025. URL: https://ceur-ws.org/Vol-4091/paper34.pdf.

[11] R. Reiter, A theory of diagnosis from first principles, Artificial intelligence 32 (1987) 57–95. doi:10.1016/0004-3702(87)90062-2.

[12] F. Wotawa, A variant of Reiter's hitting-set algorithm, Inf. Process. Lett. 79 (2001) 45–51. doi:10.1016/S0020-0190(00)00166-6.

[13] I. Pill, T. Quaritsch, RC-Tree: A variant avoiding all the redundancy in Reiter's minimal hitting set algorithm, in: 2015 IEEE International Symposium on Software Reliability Engineering Workshops, ISSRE Workshops, Gaithersburg, MD, USA, November 2-5, 2015, IEEE Computer Society, 2015, pp. 78–84. doi:10.1109/ISSREW.2015.7392050.

[14] U. Junker, QuickXplain: Preferred explanations and relaxations for over-constrained problems, in: Proceedings of the Nineteenth National Conference on Artificial Intelligence, Sixteenth Conference on Innovative Applications of Artificial Intelligence, July 25-29, 2004, San Jose, California, USA, AAAI Press / The MIT Press, 2004, pp. 167–172. URL: http://www.aaai.org/Library/AAAI/2004/aaai04-027.php.

[15] K. M. Shchekotykhin, D. Jannach, T. Schmitz, MergeXplain: Fast computation of multiple conflicts for diagnosis, in: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015, AAAI Press, 2015, pp. 3221–3228. URL: http://ijcai.org/Abstract/15/454.

[16] M. Homola, J. Pukancová, J. Boborová, I. Balintová, Merge, explain, iterate: A combination of MHS and MXP in an ABox abduction solver, in: Logics in Artificial Intelligence - 18th European Conference, JELIA 2023, Dresden, Germany, September 20-22, 2023, Proceedings, volume 14281 of *Lecture Notes in Computer Science*, Springer, 2023, pp. 338–352. doi:10.1007/978-3-031-43619-2_24.

[17] J. Boborová, J. Kloc, M. Homola, J. Pukancová, On the way to diverse datasets for evaluating abox abduction algorithms (Extended abstract), in: L. Tendera, Y. Ibáñez-García, P. Koopmann (Eds.), Proceedings of the 38th International Workshop on Description Logics - DL 2025, Opole, Poland, September 3-6, 2025, volume 4091 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2025. URL: https://ceur-ws.org/Vol-4091/paper28.pdf.