Optimal Alignment of Temporal Knowledge Bases (Extended Abstract)

Oliver Fernández Gil[1,2], Fabio Patrizi[3], Giuseppe Perelli[3] and Anni-Yasmin Turhan[4]

[1]*TU Dresden, Germany*
[2]*ScaDS.AI, Germany*
[3]*La Sapienza, Italy*
[4]*Paderborn University, Germany*

Abstract

Answering temporal CQs over temporalized Description Logic knowledge bases (TKB) is a main technique to realize ontology-based situation recognition. If the collected data in such a knowledge base is inaccurate, important query answers can be missed. In this extended abstract we report on the results from our ECAI'23 paper that introduces the TKB Alignment problem. This problem asks for a given TKB, a non-entailed temporal CQ φ and a cost measure, to compute a variant of the TKB that minimally changes the TKB, such that it entails φ and is (cost-) optimal. We investigated this problem for $\mathcal{ALC}$ TKBs and temporal CQs with LTL operators and devised a solution technique to compute (cost-optimal) alignments of TKBs that extends techniques for the alignment problem for propositional LTL over finite traces.

Keywords

Temporal reasoning, query entailment, (trace) alignment

Introduction

Temporal reasoning based on Description Logics often realizes some form of situation recognition. In such applications, a complex system is observed and its general properties are modelled in the TBox, while the data collected over time on observations made over that system is represented in a sequence of ABoxes. The actual situations to be recognized are then modeled as temporal (conjunctive) queries to be answered over the temporal knowledge base (TKB) consisting of the TBox and the sequence of ABoxes. Such TKBs can be queried by temporal conjunctive queries (TCQs), which combine linear temporal logic (LTL) with conjunctive queries. Methods for answering TCQs over TKBs and testing entailment of Boolean TCQs have been intensively investigated over the last decade ([1, 2, 3]).

If data is collected from heterogeneous sources, it can be inaccurate or it could be discretized in an unsuitable way. Thus the query need not return the expected answers. Here, the crucial problem is to find a version of the TKB that admits to detect "near misses". There are mainly two approaches developed to address the problem of errors or inaccuracies in DL (T)KBs by editing the (T)KB. In case of inconsistent TKBs, ontology repairs delete ABox axioms to restore consistent versions of the TKB [4, 5]. If information is missing in the ABoxes for the query to return the expected answers, ABox abduction, i.e., adding new ABox statements has been investigated—sofar mostly in the atemporal setting [6, 7, 8]. The newly introduced TKB-alignment problem is closely related to abduction and to computing repairs of KBs, as they change a KB to either gain a desired consequence or remove an unwanted one. However, neither of the two has yet been intensively investigated for the temporalized setting and entailment of TCQs. Furthermore, TKB Alignment computes a cost-optimal solution, which is not the common setting dedicated to abduction or repairs.

In our ECAI'23 paper [9], we have introduced and investigated the new problem of (computing a) TKB Alignment. For a given TKB Γ, a temporal CQ φ that is not entailed for Γ, the TKB Alignment problem is to compute a cost-minimal modification of the sequence of ABoxes by deletions or additions of axioms or whole ABoxes so that φ becomes entailed for the modified TKB. Sofar, this problem has not been addressed in a DL setting. We developed an approach to solve instances of this new problem for TKBs written in $\mathcal{ALC}$. The problem of Trace Alignment realizes a very similar task: for a finite trace of observations and a given property specification expressed in Linear Temporal Logic (LTL), a minimal modification of the trace is produced that satisfies the specification given in form of a query φ, e.g. see [10, 11, 12]. However, in the settings investigated for trace alignment, the observations recorded in the trace are propositional, i.e., each time point of the trace represents one of finitely many possible observables, modeled as propositions. We address the TKB Alignment problem as a Trace Alignment problem in a much richer setting, where observables are now described by DL concepts and roles, and the property specification is a temporalized query using DL atoms. Furthermore, instead of satisfaction in a single structure as in classical propositional trace alignment, the open world semantics of DLs necessitates to consider entailment, i.e. reasoning w.r.t. all models.

The TKB Alignment problem and how to solve it

We consider TKBs $\Gamma = (\mathcal{T}, \Lambda)$ consisting of a TBox $\mathcal{T}$ and a finite sequence of ABoxes Λ, where statements are expressed in $\mathcal{ALC}$. The considered query language is temporal conjunctive queries (TCQs) essentially as in [13].

We generalize the verification problem of TCQ entailment (TQE) to a *synthesis* version, consisting in *minimally* modifying the ABox sequence Λ of a TKB $\Gamma = (\mathcal{T}, \Lambda)$, to obtain a TKB $\Gamma' = (\mathcal{T}, \Lambda')$, s.t. $\Gamma' \models \varphi$. Observe that if $\Gamma \models \varphi$, the problem amounts to checking TCQ entailment. To modify the ABoxes from a TKB, we consider the two kinds of ABox operations: *insertion* and *removal* of a (concept or role) assertion α. In addition to modifying its ABoxes, a TKB can be modified by adding or removing (empty) ABoxes. Each of the operations is associated with a strictly positive cost $\mathsf{c}(m) \in \mathbb{R}^+$. The *cost* function extends to finite sequences of such operations in the natural way.

Given a TKB $\Gamma = (\mathcal{T}, \Lambda)$ and a temporal conjunctive query φ, the *TKB Alignment Problem* is to find a minimal-cost sequence of operations (if any) s.t. the TKB that results from applying this sequence to Γ entails φ.

Our technique for solving the TKB Alignment Problem builds on and extends non-trivially the approach for deciding temporal query entailment over TKBs by [13] and one for solving LTL Trace Alignment for finite traces by [10]. We extend the former approach from verification to synthesis and the latter from the propositional to the DL setting, from propositional traces to TKBs, and from finite to infinite traces, since ABox sequences are commonly interpreted over infinite FOL traces.

Our solution approach consists in reducing TKB Alignment to the shortest path problem in a so-called *Minimal-instantiation Graph*, where each edge is labelled by an atomic TKB-modification and its corresponding cost. Now, in this graph every shortest path from a suitably defined *initial* node to one node from a (suitably defined) *target* set, represents an optimal solution for a modification sequence and thus to the original TKB Alignment instance. The construction of such a Minimal-instantiation Graph is based on several intermediate structures.

To obtain these structures, we proceed similarly to [13] and use the *propositional abstraction* of a Boolean TCQ φ, where every Boolean CQ in φ is treated as a proposition and the resulting φ' essentially becomes a propositional LTL formula. This facilitates the use of standard automata constructions for LTL which we extended and lifted to treat entailment (instead of satisfaction).

To compute the weights for every edge of the Minimal-instantiation Graph, requires to solve a *local*

minimization problem that returns the minimal cost for modifying a single ABox. This is the (atemporal) *KB Alignment* problem, treated in the next section. This problem can be solved in doubly exponential time and thus TKB Alignment is solvable in triply exponential time.

The overall procedure ultimately results in an effective solution approach to the TKB Alignment problem which can assess the deviation of irregular observations w.r.t. standard ones and define corrective actions to recover a standard observation.

KB Alignment and how to solve it

In this section, we consider the *knowledge base alignment problem (KB Alignment)* for the query language of Boolean combinations of (Boolean) conjunctive queries, denoted $\mathcal{B}(\textsc{bcq})$. This query language was already investigated for $\mathcal{ALC}$ KBs in [13].

Given a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and a query $\phi \in \mathcal{B}(\textsc{bcq})$, a *solution to KB Alignment* is an ABox $\mathcal{A}'$ for which the following holds:

1. $(\mathcal{T}, \mathcal{A}')$ is consistent; and

2. $(\mathcal{T}, \mathcal{A}') \models \phi$.

The problem of *KB Alignment* consists in finding a sequence of ABox-modifications such that it is a solution to KB Alignment that is cost-optimal. The approach to solve KB Alignment consists of two steps.

The first step is to compute an upper bound on the cost of optimal solutions (if such a solution exists at all), obtained from the following trivial brute-force solution. The upper bound on the ABox-modification is obtained by computing for the input TBox $\mathcal{T}$ and query ϕ, an ABox $\mathcal{A}'$ s.t. $(\mathcal{T}, \mathcal{A}')$ is consistent and $(\mathcal{T}, \mathcal{A}') \models \phi$. Then, the cost of the *trivial modification from $\mathcal{A}$ into $\mathcal{A}'$* that empties $\mathcal{A}$ and then inserts all assertions from $\mathcal{A}'$, gives an upper bound on the cost of optimal solutions.

The second step to solve KB Alignment is to iterate over all ABox-modifications of decreasing cost starting from the upper bound, and to return a modification sequence that is a solution to KB Alignment and that is cost-optimal. More precisely, a loop enumerates all ABox-modifications with cost smaller than the upper bound, that are solutions to the KB alignment and it returns one ABox modification of minimal cost.

The computation of the initial ABox $\mathcal{A}'$ obtained by the trivial modification in our algorithm uses techniques developed for solving the *query emptiness problem* in ontology-mediated query answering—a problem introduced and investigated in [14] for various DLs (including $\mathcal{ALC}$) and the query language conjunctive queries. We show correctness and termination of the resulting algorithm and we obtain that KB Alignment is solvable for $\mathcal{ALC}$ and $\mathcal{B}(\textsc{bcq})$ in *double exponential time* [9].

Conclusions and Future Work

TKB Alignment is a new variant of the finite trace alignment problem that admits richer state and property descriptions by DL knowledge bases. Our setting uses TKBs written in $\mathcal{ALC}$, CQs with LTL operators, and a cost function for the edit operations. We have shown that TKB Alignment w.r.t. temporal CQs is effectively solvable, by developing computation methods for both TKB and KB Alignment.

Interestingly, the task of TKB Alignment is closely related to computing relaxed answers of temporal CQs as follows. Given a tuple of individuals $\bar{a}$ which is not a certain answer of a given TCQ φ, solve TKB Alignment for the Boolean TCQ obtained from φ' by assigning $\bar{a}$ to the answer variables of φ. The

costs computed during TKB Alignment for φ' are then a measure of the "distance" to a certain answer of the query. Relaxed answers are then those tuples of individuals with costs below a given threshold.

Our initial investigation on TKB Alignment uses a unitary cost measure for the edit operations mostly to ease presentation, as our approach can handle other cost measures easily. In this work, we did not regard rigid symbols, which are left for future work.

Acknowledgments

References

[1] F. Baader, S. Borgwardt, P. Koopmann, V. Thost, A.-Y. Turhan, Semantic technologies for situation awareness, Künstliche Intell. 34 (2020) 543–550.

[2] A. Artale, R. Kontchakov, A. Kovtunova, V. Ryzhikov, F. Wolter, M. Zakharyaschev, First-order rewritability of ontology-mediated queries in linear temporal logic, Artif. Intell. 299 (2021) 103536.

[3] A. Soylu, M. Giese, R. Schlatte, E. Jiménez-Ruiz, E. Kharlamov, Ö. L. Özçep, C. Neuenstadt, S. Brandt, Querying industrial stream-temporal data: An ontology-based visual approach, J. Ambient Intell. Smart Environ. 9 (2017) 77–95.

[4] C. Bourgaux, P. Koopmann, A.-Y. Turhan, Ontology-mediated query answering over temporal and inconsistent data, Semantic Web 10 (2019) 475–521.

[5] M. Bienvenu, A short survey on inconsistency handling in ontology-mediated query answering, Künstliche Intell. 34 (2020) 443–451.

[6] D. Calvanese, M. Ortiz, M. Simkus, G. Stefanoni, The complexity of conjunctive query abduction in DL-Lite, in: Proceedings of the 24th International Workshop on Description Logics (DL 2011), volume 745 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2011.

[7] W. Del-Pinto, R. A. Schmidt, ABox abduction via forgetting in $\mathcal{ALC}$, in: The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, AAAI Press, 2019, pp. 2768–2775.

[8] P. Koopmann, W. Del-Pinto, S. Tourret, R. Schmidt, Signature-based abduction for expressive description logics, in: Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020), 2020. doi:https://doi.org/10.24963/kr.2020/59.

[9] O. Fernández Gil, F. Patrizi, G. Perelli, A.-Y. Turhan, Optimal alignment of temporal knowledge bases, in: K. Gal, A. Nowé, G. J. Nalepa, R. Fairstein, R. Radulescu (Eds.), ECAI 2023 - 26th European Conference on Artificial Intelligence, volume 372 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2023, pp. 708–715.

[10] G. De Giacomo, F. M. Maggi, A. Marrella, F. Patrizi, On the disruptive effectiveness of automated planning for LTL_f-based trace alignment, in: AAAI, 2017, pp. 3555–3561.

[11] M. de Leoni, F. M. Maggi, W. M. P. van der Aalst, Aligning event logs and declarative process models for conformance checking, in: BPM, 2012.

[12] M. de Leoni, F. Maggi, W. van der Aalst, An alignment-based framework to check the conformance of declarative process models and to preprocess event-log data, Inf. Syst. 47 (2015) 258–277.

[13] F. Baader, S. Borgwardt, M. Lippmann, Temporal query entailment in the description logic SHQ, J. Web Semant. 33 (2015) 71–93.

[14] F. Baader, M. Bienvenu, C. Lutz, F. Wolter, Query and predicate emptiness in ontology-based data access, J. Artif. Intell. Res. 56 (2016) 1–59.