From: **Halland, Kenneth** Hallakj@unisa.ac.za
Subject: FW: DL 2014 notification for submission 42
Date: 26 May 2014 at 14:15
To: Arina Britz arina.britz@meraka.org.za, szymon.klarman@gmail.com

---

FYI

____
From: DL 2014 [dl2014@easychair.org]
Sent: Monday, May 26, 2014 1:53 PM
To: Halland, Kenneth
Subject: DL 2014 notification for submission 42

Dear Ken,

We are pleased to inform you that your submission

42: TBox abduction in ALC using a DL tableau

has been accepted for poster presentation at the 27th International
Workshop on Description Logics (DL 2014) to be held from July 17-20 in Vienna, Austria. You will be given a brief (3min) slot to advertise your work in front of the workshop participants, in addition to the discussion at the poster board. You can display one or two slides for your plenary announcement; we will collect slides in PDF format beforehand for that purpose.

We received 80 submissions this year, which is a very high number, and the competition for full talk slots has therefore been very strong. Overall, 34 submissions have been accepted for oral presentation, and another 38 submissions have been accepted for poster presentation. This is similar to the number of accepted submissions in 2013.

Please find below some important information regarding preparation of the final version and attending the workshop (additional information can be found on the DL and Vienna Summer of Logic websites).

FINAL VERSION

You will find attached the comments of the reviewers. After carefully revising your paper, please upload the final version (PDF and the zipped LaTeX sources) into EasyChair

*** no later than June 15th, 2014. ***

This deadline is firm, and it is critical that you comply with it. No extensions will be given as the organizers need some time to prepare the workshop proceedings.

The format is the same as for the submission: full papers may be up to 11 pages in LNCS style, excluding references, while extended abstracts may be up to 3 pages, excluding references. Note that appendices should not be included in the final version, but we invite authors to include a reference or link to a longer version (this is especially encouraged in the case of extended abstracts).

REGISTRATION

At least one of the authors is required to register for the workshop. Registration information is available at

http://www.dbai.tuwien.ac.at/dl2014/registration.php

Please note that the cutoff date for early registration is:

June 8th, 2014

ATTENDING OTHER VSL EVENTS

If you are registered for a VSL event (like the DL workshop), then you can attend all scientific lectures which happen on the days of this event.

TRAVELLING TO VIENNA

There is a 15% discount for Austrian Airlines flights to Vienna (including code sharing flights) booked through the Austrian.com website. In order to use the discount, enter the following code "VSL14" in the eVoucher field.

We are looking forward to seeing you in Vienna at DL 2014!

Best regards,

Meghyn and Riccardo
DL 2014 PC Chairs

----------------------- REVIEW 1 ---------------------
PAPER: 42
TITLE: TBox abduction in ALC using a DL tableau
AUTHORS: Ken Halland, Arina Britz and Szymon Klarman

OVERALL EVALUATION: 1 (weak accept)
REVIEWER'S CONFIDENCE: 4 (high)

----------- REVIEW -----------
This submission studies the problem of finding explanations for missing entailments in ALC TBoxes. Solutions to the problem consist in a set of GCIs which, together with the given TBox, support the entailment. This problem is relevant for repairing ontologies that lack expected entailments. The main contribution is an algorithm for finding solutions in a restricted language L_min (allowing only atomic concepts, negations thereof, and two kinds of unqualified restrictions). This algorithm runs the standard tableau algorithm, visits all open branches, identifies all sets of L_min axioms required to close them, and runs Reiter's Hitting Set Tree algorithm on those sets. The submission contains illustrating examples, proofs for soundness and completeness of the algorithm, a complexity analysis and a description of how to handle two generalizations to the initial problem.

Clearly this is interesting work because the abduction problem has received less attention in the literature than other, equally relevant, forms of explanations such as axiom pinpointing aka finding justifications for positive entailments. The general idea of "recruiting" solutions from statements that close all open tableau branches is elegant, but it has already been used in previous work since the 1990's, including in the authors' recent work on ABox abduction. The novelty of the current approach is to find a solution a specific variant of the abduction problem which does not seem to have been treated before (solutions are GCIs), and to "under-approximate" the set of solutions in an efficient way. It is somewhat unfortunate that this "under-approximation" is rather straightforward and provides a quite restrictive language for formulating solutions, which is owed to the requirement of efficiency.
This apparently strong restriction and its implications should be discussed in some depth, in my opinion. Furthermore, the presentation suffers from some technical flaws (which are all repairable as far as I can see), and some explanations should be more precise; details follow. Altogether, the approach deserves publication, but since its novelty is limited, it might better be presented as a poster.

Technical flaws
----

* Semantically minimal solutions do not have to exist, neither in ALC nor in L_min. The reader may already ask herself this question in Section 3, but the issue is not addressed before the end of Section 4. That remark is rather weak: it should give some non-trivial example that does not have s.m. solutions in ALC, and one that has s.m. solutions in ALC but not in L_min. Then it should be discussed how serious both restrictions are expected to be for the real-world problem that is modelled here.

* Internalization applies to all tableau nodes, not only to the initial node. This was overlooked, affecting Example 2 and the proof of Theorem 1, where it causes more branching than currently acknowledged, and thus more solutions too. More precisely: in Example 2, there are two open branches in addition to the closed one, because the internalized axiom of K is added to the R-successor d of c. The first open branch contains {A1(c), R(c,d), A2(d), -A3(d), -A1(d)} and leads to the new solution candidates A2 [= A1 and -A1 [= A3. The second open branch contains {A1(c), R(c,d), A2(d), -A3(d), R(d,e), A2(e)} (e is blocked by d) and leads to the new solution candidates A2 [= forall R.bot, -A3 [= forall R.bot. Each combination of these new solution candidates leads to a new solution that has not been considered yet. Furthermore, in the proof sketch and proof of Theorem 1, one needs to acknowledge that, via internalization, the tableau T' for K+Theta has significantly more branches t!

han the tableau T for Theta. For the sake of a precise argument, one might need to establish a mapping from branches in T' to branches in T and argue that all branches in T' contain their image. This implies that the branches whose image is closed are closed too, and the remaining branches can be treated as in the current proof.

* In Example 1, the solution A1 [= -A2 is discareded because it allegedly contradicts the original knowledge base. Strictly speaking, it does not -- it only makes A1 unsatisfiable if added to K, which does not contradict the consistency requirement as it is stated now. The same is the case for the axiom exists R.top [= -A1 in Example 2. Either these axioms are allowed as solutions, or incoherency needs to be ruled out along with inconsistency, which would then rule out the "interesting" solutions exists R. [= forall R.bot in Example 2 (which makes role R unsatisfiable) and -A1 [= A2 in Example 2 (which makes A2 equivalent to top, which is often considered as dubious as a concept being unsatisfiable).

* The complexity analysis is problematic in three points: (1) the announcement "we can do better than this" at the top of page 10 is not fulfilled because the described optimization of the algorithm still runs in doubly exponential time in the worst case (all branches may have to be visited). (2) The point about Reiter's algorithm is unclear: if its space requirements are exponential in its input as it is said here, then its application requires doubly exponential space w.r.t. the input ontology. The only way to remain within EXPSPACE is to analyze Reiter's algorithm more carefully and conclude that it requires space polynomial in the _output_. Together with the pointer to the exponential number of solutions, the argument would be complete. (3) The postprocessing needs to be considered in the complexity analysis too.

Requests for more precision
----

* Page 6, why are solution candidates L [= -L discarded? At first sight, they seem neither irrelevant nor inconsistent. It seems to have to do with Definition 1 requiring C,D to be satisfiable, but this should be explained more carefully.

* The structure of the proof of Theorem 2 is confusing because the case distinction cannot easily be seen to be complete: for example, it

does not explicitly rule out the case that Theta contains two simple concept inclusions that do not form a chain, such as L1 [= L2, L3 [= L4. Instead, one should argue what would happen if Theta contained more than one axiom. Fix two axioms in Theta and consider cases regarding the forms of these axioms. Many of these 16 cases can be argued away in one sentence that is currently repeated in (a)-(c). The remaining cases are essentially (a)-(c) minus the redundant statements.

Further comments
----

* Intro, line 1: the use cases that are mentioned apply to ABox abduction only. TBox abduction is used for debug or repair.

* Intro, explanation of "relevance": what does "introduce an independent theory" mean? It may be read as if the solution should use the same vocabulary as the TBox, which is not the definition of relevance. One of the important points about relevance, in my understanding, is that it prohibits the entailment itself to be a solution.

* Proposition 1, Proof (a), second sentence: it has to be said that Theta' is a solution.

* Proposition 1, (b) is really trivial and shouldn't require a separate proof.

Further questions, out of interest
----

* How is this approach expected fare in practice? After all, the runtime is still doubly exponential, and the current formulation seems to prohibit the use of a tableau reasoner as a black box (however, this might be possible with a more refined variant of the current algorithm). Furthermore, it would be interesting to know the "success rate" for real-world abduction instances -- i.e., in how many cases does a semantically minimal solution in ALC / L_min exist?

* Is it realistic to expect significant results from an extension of this approach to more expressive DLs when this rather inexpressive language L_min is maintained?

----------------------- REVIEW 2 ---------------------
PAPER: 42
TITLE: TBox abduction in ALC using a DL tableau
AUTHORS: Ken Halland, Arina Britz and Szymon Klarman

OVERALL EVALUATION: 1 (weak accept)
REVIEWER'S CONFIDENCE: 4 (high)

----------- REVIEW -----------
The paper addresses the problem of TBox abduction in ALC. Major contributions include a definition of TBox abduction and a tableau-based algorithm for computing abductive solutions (the idea is straightforward). It is proven that the proposed algorithm is sound for ALC but it is only complete for ALC_min, a very restricted fragment of ALC. While some possible optimisations are briefly discussed, no implementation is reported in the paper. The paper is well written in general.

Detailed comments:
Given that the proposed algorithm is incomplete for ALC, it would be necessary to provide a counter example for the incompleteness. E.g., K = {hasParent(a,b), \exists hasParent.Person \sqsubseteq Person} and the observation is Person(a). Obviously, K does not entail the observation and an abductive solution is { \exists hasParent \sqsubseteq Person}. But the algorithm could not find this solution because the open branch does not fall into any option in step 2 of the Algorithm.

----------------------- REVIEW 3 ---------------------
PAPER: 42
TITLE: TBox abduction in ALC using a DL tableau
AUTHORS: Ken Halland, Arina Britz and Szymon Klarman

OVERALL EVALUATION: -1 (weak reject)
REVIEWER'S CONFIDENCE: 4 (high)

----------- REVIEW -----------
The paper presents a tableau-based algorithm for TBox abduction in ALC that produces explanations expressed only within a restricted language.

The paper is well-written and structured. My major concern is its technical contributions, novelty and new material presented, as the problem and solution to abduction is more-or-less pretty standard and known. So you first saturate using the tableau rules and then from the open branches of the constructed tableau you attempt to figure out axioms that if were present in the TBox these branches would close. This idea is already known in first-order logic and has been applied before even in DLs. To name a few papers:

Tommaso Di Noia, Eugenio Di Sciascio and Francesco M. Donini. A Tableaux-based calculus for Abduction in Expressive Description Logics: Preliminary Results. 22nd International Workshop on Description Logics (DL2009)

S. Colucci , T. Di Noia , E. Di Sciascio , F. M. Donini , M. Mongiello. A Uniform Tableaux-Based Approach to Concept Abduction and Contraction in ALN. In Proc. of DL-04 (2004). CEUR Workshop

(2004), CLOR Workshop.

Marta Cialdea Mayer , Fiora Pirri. First Order Abduction Via Tableau and Sequent Calculi.
Bulletin of the IGPL.

Perhaps the abduction problem is not the same as the one studied in the above papers (I haven't checked), however, the above idea of building axioms from the open branches is more-or-less typical.

As a conclusion it is not clear what is the new techniques and the problems addressed in the current paper.

I believe that the notion of internalisation is not used properly. Internalisation is the process of converting the whole TBox into a single concept assertion and then in resoning checking satisfiability of that single concept without ever looking at the TBox again. I don't think this is possible in ALC. What is mentioned at the end of page 2 \bigpsi should be done for each fresh/new individual that is introduced by the \exists-rule. So the algorithm needs to check the TBox as it goes along. So this is not realy internalisation.

The paper would definately improve if an implementation and experimental evaluation is also performed.

In page 5 you repeat the same information twice, that is, steps 1--3 at the beginning of the page and steps 1--3 at the end. I believe that the first is not needed.

In page 7: "But together with the observation, it makes A_1 equivalent" Do you rather mean "together with the TBox".

It would be better if you used a different name for the problem defined in definition 1 from the one defined in definition 2 something like generalised abductive solution/problem. A second problem with this is that then in theorems 1 and 2 it is not clear to which problem you are really referring.

top of page 8: the negation of the such -> the negation of such.