A Uniform Tableaux-Based Method for Concept Abduction and Contraction in Description Logics

S. Colucci and **T. Di Noia** and **E. Di Sciascio** and **F. M. Donini *** and **M. Mongiello** [1]

1 The Calculus

We present algorithms for Concept Abduction and Concept Contraction, two reasoning services in Description Logics (DL) recently proposed to model how several supplies fit a demand (all described by concepts), and vice versa, in e-commerce. An extended version of the paper is in [2]. Recent papers on tableaux for description logics use a labeling function $\mathcal{L}$ to map an individual$x$ to a set of concepts $\mathcal{L}(x)$ such that for every concept $C, C \in \mathcal{L}(x)$ stands for the formula $C(x)$, and similarly for roles $R \in \mathcal{L}(x,y)$. Here we distinguish between formulas labeled "true" and formulas labeled "false" in the tableaux, hence we use two labeling functions **T**() and **F**(), both going from individuals to sets ofconcepts, and from pairs of individuals to sets of roles. A (usual) tableau branch is now represented by two functions **T**() and **F**(). Moreover, we write in the name of an individual x its history, *i.e.*, the string identifying x is made up of integers and role symbols, such as $x = 1R3Q7$, which means that individual x is used for concepts in a quantification involving role R, and inside, a quantification involving role Q. Integers in between roles make sure that such strings are unique, *i.e.*, there can be two individuals with the same role sequence, but not with the same integer sequence [3]. Given an individual x in a tableau, an interpretation $(\Delta^\mathcal{I}, \cdot^\mathcal{I})$ satisfies two tableau labels $\mathbf{T}(x)$ and $\mathbf{F}(x)$ if, for every concept $C \in \mathbf{T}(x)$ and every concept $D \in \mathbf{F}(x)$, it is $x^\mathcal{I} \in C^\mathcal{I}$ and $x^\mathcal{I} \notin D^\mathcal{I}$ respectively. Similarly, $(\Delta^\mathcal{I}, \cdot^\mathcal{I})$ satisfies two tableau labels $\mathbf{T}(x,y)$ and $\mathbf{F}(x,y)$ if for every role $R \in \mathbf{T}(x,y)$ and for every role $Q \in \mathbf{F}(x,y)$ it holds $(x^\mathcal{I}, y^\mathcal{I}) \in R^\mathcal{I}$ and $(x^\mathcal{I}, y^\mathcal{I}) \notin Q^\mathcal{I}$. We note however that for the DL we adopt, every role Q that appears in a label $\mathbf{F}(x,y)$ is of the form $\neg R$, hence $Q \in \mathbf{F}(x,y)$ means, in fact, $(x^\mathcal{I}, y^\mathcal{I}) \in R^\mathcal{I}$ too. An interpretation satisfies a tableau branch if it satisfies $\mathbf{T}(x)$, $\mathbf{F}(x)$, $\mathbf{T}(x,y)$ and $\mathbf{F}(x,y)$ for every individual x, and for every pair of individuals x, y in the branch. Each rule has a precondition, and an action modifying the tableau. When the precondition is met, the action can be performed. In order to simplify the preconditions, we assume that, for each different instance of the preconditions, each rule is applied at most once. We also assume that concepts are always simplified in Negation Normal Form (NNF, see [1, ch.2]), so that negations come only in front of concept names. Without NNF, the number of rules should be doubled. In what follows, given a concept C, we denote with $\overline{C}$ the NNF of $\neg C$. Rules come in pairs, first the (usual) version with a construct in the **T**-constraints, then the dual construct in the **F**-constraints. However, in what follows we omit rules marked with an asterisk (*), because the correspondent formulae do not appear in our tableaux for $\mathcal{ALN}$.

1. conjunctions:

T⊓) if $C \sqcap D \in \mathbf{T}(x)$, then add both C and D to $\mathbf{T}(x)$

F⊔) if $C \sqcup D \in \mathbf{F}(x)$, then add both C and D to $\mathbf{F}(x)$

2. disjunctions (branching rules):

T⊔) *

F⊓) if $C \sqcap D \in \mathbf{F}(x)$, then add either C or D to $\mathbf{F}(x)$

3. existential quantifications:

T∃) *

F∀) if $\forall R.C \in \mathbf{F}(x)$, then pick up a new individual $y = x \circ R \circ m$ (where m is an integer such that y is unique), add $\neg R$ to $\mathbf{F}(x,y)$, and let $\mathbf{F}(y) := \{C\}$

4. universal quantifications:

T∀) if $\forall R.C \in \mathbf{T}(x)$ and there exists an individual y such that either $R \in \mathbf{T}(x,y)$, or $\neg R \in \mathbf{F}(x,y)$, then add C to $\mathbf{T}(y)$

F∃) if $\exists R.C \in \mathbf{F}(x)$, and there exists an individual y such that either $R \in \mathbf{T}(x,y)$, or $\neg R \in \mathbf{F}(x,y)$, then add C to $\mathbf{F}(y)$

5. at-least number restrictions:

T⩾) if $\geqslant n\, R \in \mathbf{T}(x)$, with $n > 0$, and for every individual y neither $R \in \mathbf{T}(x,y)$ nor $\neg R \in \mathbf{F}(x,y)$, then pick up a new individual $y = x \circ R \circ m$ (where m is an integer such that y is unique), add R to $\mathbf{T}(x,y)$, and let $\mathbf{T}(y) := \emptyset$

F⩽) if $\leqslant n\, R \in \mathbf{F}(x)$ and for every individual y neither $R \in \mathbf{T}(x,y)$ nor $\neg R \in \mathbf{F}(x,y)$, then pick up a new individual $y = x \circ R \circ m$ (where m is an integer such that y is unique), add $\neg R$ to $\mathbf{F}(x,y)$, and let $\mathbf{F}(y) := \emptyset$

6. at-most number restrictions:

T⩽) if $\leqslant 1\, R \in \mathbf{T}(x)$, and there are 2 individuals y_1, y_2 such that for $i \in 1,2$ it is either $R \in \mathbf{T}(x, y_i)$ or $\neg R \in \mathbf{F}(x, y_i)$, then let $\mathbf{T}(y_1) := \mathbf{T}(y_1) \cup \mathbf{T}(y_2)$, let $\mathbf{F}(y_1) := \mathbf{F}(y_1) \cup \mathbf{F}(y_2)$, and eliminate y_2 in the branch

F⩾) if $\geqslant 2\, R \in \mathbf{F}(x)$ and there are 2 individuals y_1, y_2 such that for $i \in 1,2$ it is either $R \in \mathbf{T}(x, y_i)$ or $\neg R \in \mathbf{F}(x, y_i)$, then let $\mathbf{T}(y_1) := \mathbf{T}(y_1) \cup \mathbf{T}(y_2)$, let $\mathbf{F}(y_1) := \mathbf{F}(y_1) \cup \mathbf{F}(y_2)$, and eliminate y_2 in the branch

7. axioms in $\mathcal{T}$:

F⊑) if x is an individual, and $A \sqsubseteq C \in \mathcal{T}$, then add $A \sqcap \overline{C}$ to $\mathbf{F}(x)$

F≐) if is an individual, and $A \doteq C \in \mathcal{T}$, then add both $A \sqcap \overline{C}$ and $C \sqcap \neg A$ to $\mathbf{F}(x)$

When more than one rule can be applied, we always give *lowest* precedence to Rules **T⩾)** and **F⩽)**, while other rules can be applied

[1] Politecnico di Bari, Via Re David, 200, I-70125, Bari, Italy * Università della Tuscia, via San Carlo, 32I, -01100, Viterbo, Italy

in any order. We now split the definition of clash (an explicit inconsistency) between clashes involving the same truth prefix (homogeneous clashes) and those involving both prefixes (heterogeneous clashes).

Definition 1 (Clash) *A branch contains a* homogeneous clash *if it contains one of the following:*

1. *either $\bot \in \boldsymbol{T}(x)$ or $\top \in \boldsymbol{F}(x)$, for some individual x;*
2. *either $A, \neg A \in \boldsymbol{T}(x)$ or $A, \neg A \in \boldsymbol{F}(x)$ for some individual x and some concept name A;*
3. *either $\geqslant n\,R, \leqslant m\,R \in \boldsymbol{T}(x)$ with $m < n$, or $\leqslant n\,R, \geqslant m\,R \in \boldsymbol{F}(x)$ with $m - 1 < n + 1$, for some individual x, and some role name R.*

A branch contains a heterogeneous clash *if it contains one of the following:*

1. *$\boldsymbol{T}(x) \cap \boldsymbol{F}(x)$ contains either A or $\neg A$ for some individual x and some concept name A;*
2. *either $\geqslant n\,R \in \boldsymbol{T}(x)$ and $\geqslant m\,R \in \boldsymbol{F}(x)$ with $m - 1 < n$, or $\leqslant n\,R \in \boldsymbol{T}(x)$ and $\leqslant m\,R \in \boldsymbol{F}(x)$ with $n < m + 1$, for some individual x, and some role R*

2 Algorithms

We now present the two algorithms for Concept Contraction and Concept Abduction. We need to define a function, $roles(x)$, that given an individual x (as a sequence of integers and roles) returns the sequence of roles in x (without integers). For example, $roles(1R3Q7) = RQ$. Moreover, we denote the substitution of an occurrence of a concept C with the concept $\top$, inside a concept D as $D[C \to \top]$.

Algorithm $contract$
input: $\mathcal{ALN}$ concepts C, D, acyclic TBox $\mathcal{T}$
output: concepts K (keep), G (giveup)
begin
  compute a complete tableau τ for $\mathcal{T}$, $D \in \mathbf{T}(x), \overline{C} \in \mathbf{F}(x)$
  IF τ is open **THEN**
    RETURN $G := \top, K := D$ /* no contraction needed */
  ELSE IF every branch in τ contains a homogeneous clash **THEN**
    RETURN fail /* either C or D is unsatisfiable in $\mathcal{T}$ */
  ELSE
    choose(*) a branch β containing only heterogeneous clashes;
    LET $\mathcal{G} := \{\langle C_i, x_i\rangle | C_i \in \mathbf{T}(x_i), \overline{C_i} \in \mathbf{F}(x_i)$ is a clash in $\beta\}$;
    LET $G := \sqcap_{\langle C_i, x_i\rangle \in \mathcal{G}} \forall roles(x_i).C_i$
    LET $K := D[C_i \to \top]_{\langle C_i, x_i\rangle \in \mathcal{G}}$
    RETURN G, K
END

Observe that the algorithm admits a nondeterministic choice in step (*), needed to select the contraction according to some minimality criterion, and that only branches without homogeneous clashes need to be completely expanded, even after the first clash has been found. Observe also that substituting an occurrence of a concept C with $\top$ corresponds, in $\mathcal{ALN}$, to eliminating the occurrence. In fact, since $\top \sqcap D = D$, and $\forall R.\top = \top$, once $\top$ has been inserted, the concept simplifies eliminating it from conjunctions and universal quantifications. We preferred this notation instead of eliminating occurrences, since it seems us to be more concise.

Theorem 1 *The concepts G, K returned by the Algorithm $contract$ are a Contraction of D w.r.t. C and $\mathcal{T}$.*

Proof. First, note that $K \sqcap C$ is satisfiable by definition of K; in fact, the tableau for $K \sqcap C$ is the same as the tableau for $D \sqcap C$, but it has now at least one open branch β, in which all clashes have been removed. Secondly, $D = G \sqcap K$ by construction. □

Note that Algorithm $contract$ proves that Concept Contraction in $\mathcal{ALN}$ with bushy TBoxes is solvable in polynomial time. We now present the algorithm for Concept Abduction. Also this algorithm uses the tableaux rules defined in Section 1.

Algorithm $abduce$
input: $\mathcal{ALN}$ concepts C, D, acyclic TBox $\mathcal{T}$
output: concept H (hypotheses)
begin
  compute a complete tableau τ for $\mathcal{T}$, $C \in \mathbf{T}(x), D \in \mathbf{F}(x)$
  IF τ is closed **THEN**
    RETURN $H := \top$ /* no abduction needed */
  ELSE
    choose(*) a set of pairs $\mathcal{H} := \{\langle C_i, x_i\rangle\}$ and
    LET $H := \sqcap_{\langle C_i, x_i\rangle \in \mathcal{H}} \forall roles(x_i).C_i$
    such that (1) every open branch in τ contains at least one
      constraint $C_i \in \mathbf{F}(x_i)$ from $\mathcal{H}$
      (2) $C \sqcap H$ is satisfiable in $\mathcal{T}$
    RETURN H
END

Theorem 2 *The concept H returned by the Algorithm $abduce$ is a solution of the CAP $\langle C, D, \mathcal{T}\rangle$.*

Proof. First, we have $\mathcal{T} \models C \sqcap H \sqsubseteq D$, since the tableau starting from $C \sqcap H \in \mathbf{T}(1), D \in \mathbf{F}(1)$ is τ, plus the constraints signed **T** from H. These constraints include $C_i \in \mathbf{T}(x_i)$, making every open branch in τ closed by a heterogeneous clash. Regarding the condition $C \sqcap H$ satisfiable in $\mathcal{T}$, it is enforced by Condition (2) in the choice of $\mathcal{H}$. □

Condition (2) is necessary in $abduce$, since heterogeneous clashes could be formed also by contradicting an axiom in $\mathcal{T}$. In that case, although still $C \sqcap H \sqsubseteq D$ in $\mathcal{T}$, the subsumption trivially holds since $C \sqcap H = \bot$. We conclude the section by showing that our Algorithm $abduce$ puts an upper bound to Concept Abduction that meets the lower bound proved in the previous section.

Theorem 3 *Let $\mathcal{P} = \langle C, D, \mathcal{T}\rangle$ a Concept Abduction Problem, where C, D are concepts in $\mathcal{ALN}$, and $\mathcal{T}$ is a bushy TBox in $\mathcal{ALN}$. Deciding whether there exists a solution of length k in $SOLCAP_{\leq}(\mathcal{P})$ is NP-complete.*

The proof of the theorem is in [2].

Acknowledgements

We thank Andrea Calì and Diego Calvanese for useful discussions about DL. We also acknowledge the project PON CNOSSO and Italian MIUR under project PRIN 2002 for supporting this research.

REFERENCES

[1] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel Schneider (eds.), *The Description Logic Handbook*, Cambridge University Press, 2003.

[2] S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello, 'Uniform Tableaux-Based Approach to Concept Abduction and Contraction in ALN DL', in *Proc. of the 17th Intl. Workshop on Description Logics (DL'04)*, CEUR Workshop Proc., (2004). To appear.

[3] F.M. Donini and F. Massacci, 'Exptime tableaux for $\mathcal{ALC}$', *Artif. Intell.*, **124**, 87–138, (2000).