Fachbereich Informatik der Universität Hamburg

Vogt-Kölln-Str. 30 ◊ D-22527 Hamburg / Germany

University of Hamburg - Computer Science Department

Mitteilung Nr. 288/99 • Memo No. 288/99

Expressive ABox Reasoning with Number Restrictions, Role Hierarchies, and Transitively Closed Roles

(Revised Version)

Volker Haarslev and Ralf Möller

Arbeitsbereich KOGS

FBI-HH-M-288/99

January 2000

Expressive ABox Reasoning with Number Restrictions, Role Hierarchies, and Transitively Closed Roles

Volker Haarslev and **Ralf Möller**
University of Hamburg, Computer Science Department,
Vogt-Kölln-Str. 30, 22527 Hamburg, Germany
`http://kogs-www.informatik.uni-hamburg.de/~<name>/`

Abstract

We present a new tableaux calculus deciding the ABox consistency problem for the expressive description logic $\mathcal{ALCNH}_{R^+}$. Prominent language features of $\mathcal{ALCNH}_{R^+}$ are number restrictions, role hierarchies, transitively closed roles, and generalized concept inclusions. The ABox description logic system RACE [Haarslev et al., 1999] is based on the calculus for $\mathcal{ALCNH}_{R^+}$.

A revised version of this report appears in: *Proceedings of Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR2000), Fausto Giunchiglia and Bart Selman (eds), Breckenridge, Colorado, USA, 12-15 April 2000.*

1 Introduction

Experiences with concept languages indicate that at least description logics (DLs) with negation and disjunction are required to solve practical modeling problems without resorting to ad hoc extensions. The requirements derived from practical applications of DLs ask for even more expressive languages. For instance, in [Sattler, 1996] the need for transitive roles is demonstrated for representing part-whole relations, family relations or partial orders in general. It is argued that the trade-off between expressivity and complexity favors the integration of transitively closed roles instead of a transitive closure operator for roles. Other examples are given in [Horrocks, 1998], where the area of medical terminology is discussed. Design studies for the Galen project identified the need for modeling of transitive part-whole, causal and compositional relations, and to organize these relations into a hierarchy.

Moreover, generalized concept inclusions were also required as a modeling tool, e.g. for expressing sufficient conditions of concepts.

2 The Description Logic $\mathcal{ALCNH}_{R^+}$

Motivated by the above-mentioned requirements we introduce in this paper an ABox tableaux calculus for the description logic $\mathcal{ALCNH}_{R^+}$. It augments the basic logic $\mathcal{ALC}$ [Schmidt-Schauss and Smolka, 1991] with number restrictions, role hierarchies, and transitively closed roles. Note that these language features imply the presence of generalized concept inclusions and cyclic concepts. The use of number restrictions in combination with transitive roles and role hierarchies is syntactically restricted: no number restrictions are possible for (i) transitive roles and (ii) for any role which has a transitive subrole. Furthermore, we assume that the unique name assumption holds for ABox individuals.

$\mathcal{ALCNH}_{R^+}$ is an extension of $\mathcal{ALCNH}$ that itself can be polynomially reduced to $\mathcal{ALCNR}$ [Buchheit et al., 1993] and vice versa. It is possible to rephrase every hierarchy of role names with a set of role conjunctions and vice versa [Buchheit et al., 1993]. Thus, our work on $\mathcal{ALCNH}_{R^+}$ extends the work on $\mathcal{ALCNR}$ by additionally providing transitively closed roles. $\mathcal{ALCNH}_{R^+}$ also extends other related description logics such as $\mathcal{ALC}_{R^+}$ [Sattler, 1996] and $\mathcal{ALCH}f_{R^+}$ [Horrocks, 1998]. The work on these logics has been extended and a tableaux calculus for deciding concept consistency for the language $\mathcal{ALCQHI}_{R^+}$ has been presented in [Horrocks et al., 1999b]. Another approach is presented in [De Giacomo and Lenzerini, 1996] where the logic $\mathcal{CIQ}$ for reasoning with TBoxes and ABoxes is introduced. The reasoning procedures developed for $\mathcal{CIQ}$ are based on a polynomial encoding of $\mathcal{CIQ}$ TBoxes into sublanguages of $\mathcal{CIQ}$. A similar approach is taken for ABoxes of the languages $\mathcal{CI}$ and $\mathcal{CQ}$. In comparison to $\mathcal{ALCNH}_{R^+}$ and the other approaches mentioned above $\mathcal{CIQ}$ offers more operators (e.g. the transitive closure) but does not support role hierarchies and allows number restrictions only for primitive roles.

2.1 ABox Reasoning versus Concept Consistency

$\mathcal{ALCNH}_{R^+}$ can be considered as a sublanguage of $\mathcal{ALCQHI}_{R^+}$ if we neglect the ABox part of $\mathcal{ALCNH}_{R^+}$. However, ABox reasoning truly extends the usefulness of description logics in practical applications. The increase of expressiveness is also reflected in an increase of the complexity of the tableaux rules (see Section 4.1 for more details). An alternative might be the so-called "precompletion approach" originally developed for the language $\mathcal{ALCQ}$ [Hollunder, 1994] and recently adapted to $\mathcal{ALCH}_{R^+}$ [Tessaris and Gough, 1999].

The idea behind the precompletion approach is to transform given ABoxes in a way such that ABox satisfiability is reduced to concept satisfiability. This is achieved by generating a precompletion of an ABox where all role filler relationships between ABox individuals (e.g. $\mathsf{i}_1, \ldots, \mathsf{i}_n$) have been absorbed into corresponding concept terms (e.g. $\mathsf{C}_1, \ldots, \mathsf{C}_n$). Then, ABox consistency can be reduced to testing the satisfiability of a concept conjunction (e.g. $\mathsf{C}_1 \sqcap \ldots \sqcap \mathsf{C}_n$). The advantage of this approach is that it allows one to reuse existing tableaux provers for concept consistency. However, there currently exist no calculi for computing the precompletion of ABoxes for languages such as $\mathcal{ALCNH}_{R^+}$ or even $\mathcal{ALCQHI}_{R^+}$. Moreover, for practical applications one can argue that a translational approach via precompletion techniques might raise problems for relating results from the concept consistency tester (e.g. concept incoherence) to corresponding ABox individuals and their assertions (e.g. ABox incoherence).

Another difficulty results in the applicability of optimization techniques such as dependency-directed backtracking (e.g. see [Horrocks and Patel-Schneider, 1999] for a discussion of these techniques). The translational approach will need similar information from a concept consistency tester in order to avoid unnecessary backtracking. An ABox tableaux calculus can avoid this problem as illustrated with the following ABox $\mathcal{A}$.

$$\mathcal{A} := \{\mathsf{i} : \forall \mathsf{R} . \mathsf{A} \sqcup \forall \mathsf{R} . \mathsf{B},\ (\mathsf{i}, \mathsf{k}) : \mathsf{R},\ (\mathsf{j}, \mathsf{k}) : \mathsf{R},\ \mathsf{j} : \mathsf{C} \sqcup \mathsf{D},\ \mathsf{k} : \neg \mathsf{A} \sqcap \neg \mathsf{B}\}$$

Our ABox tableaux calculus (see below for details) non-deterministically generates new ABoxes for dealing with disjunctions. An effective search procedure has to use backtracking to exhaustively explore all possible alternatives. For instance, the case $\mathsf{i} : \forall \mathsf{R} . \mathsf{A}$ must be explored. Another choice point is the disjunction $\mathsf{j} : \mathsf{C} \sqcup \mathsf{D}$. Without loss of generality, the system tries $\mathsf{j} : \mathsf{C}$ before considering constraints for k (note that C might contain concept value restrictions). Afterwards, the concept constraint $\mathsf{i} : \forall \mathsf{R} . \mathsf{A}$ is treated in combination with the role constraint $(\mathsf{i}, \mathsf{k}) : \mathsf{R}$. After some additional expansion steps, this part of the search tree will lead to a clash because $\mathsf{k} : \mathsf{A}$ and $\mathsf{k} : \neg \mathsf{A}$ will be elements of the ABox. Now, if the system backtracks to the choice point $\mathsf{j} : \mathsf{C} \sqcup \mathsf{D}$ and tries $\mathsf{j} : \mathsf{D}$ it is bound to detect the same clash for k again. Since D can be a very complex concept term, many expansion steps are definitely waisted. Using dependency-directed backtracking, one detects that $\mathsf{j} : \mathsf{C}$ is not involved in the clash and backtracking is set up again for trying $\mathsf{i} : \forall \mathsf{R} . \mathsf{B}$.

2.2 The Concept Language

We present the syntax and semantics of the language for specifying concept and role inclusions.

Definition 1 (Role Inclusions, Role Hierarchy) Let P and T be disjoint sets of non-transitive and transitive *role* names, respectively, and let R be defined as $R = P \cup T$. Let R and S be role names, then $\mathsf{R} \sqsubseteq \mathsf{S}$ *(role inclusion axiom)* is a terminological axiom. Given a set of role inclusion axioms, we define a *role hierarchy* where $\sqsubseteq^*$ is the reflexive transitive closure of $\sqsubseteq$ over R.

Additionally we define the set of ancestors and descendants of a role.

Definition 2 (Role Descendants/Ancestors) Given a role hierarchy the set $\mathsf{R}^{\uparrow} := \{\mathsf{S} \in R \,|\, \mathsf{R} \sqsubseteq^* \mathsf{S}\}$ defines the *ancestors* and $\mathsf{R}^{\downarrow} := \{\mathsf{S} \in R \,|\, \mathsf{S} \sqsubseteq^* \mathsf{R}\}$ the *descendants* of a role R. We also define the set $S := \{\mathsf{R} \in P \,|\, \mathsf{R}^{\downarrow} \cap T = \emptyset\}$ of *simple* roles that are neither transitive nor have a transitive role as descendant.

Definition 3 (Concept Terms) Let C be a set of concept names which is disjoint from R. Any element of C is a *concept term*. If C and D are concept terms, $\mathsf{R} \in R$ is an arbitrary role, $\mathsf{S} \in S$ is a simple role, $n > 1$, and $m > 0$, then the following expressions are also concept terms:

- $\top$ *(top concept)*
- $\bot$ *(bottom concept)*
- $\mathsf{C} \sqcap \mathsf{D}$ *(conjunction)*
- $\mathsf{C} \sqcup \mathsf{D}$ *(disjunction)*
- $\neg \mathsf{C}$ *(negation)*
- $\forall \mathsf{R} \,.\, \mathsf{C}$ *(concept value restriction)*
- $\exists \mathsf{R} \,.\, \mathsf{C}$ *(concept exists restriction)*
- $\exists_{\leq m} \mathsf{S}$ *(at most number restriction)*
- $\exists_{\geq n} \mathsf{S}$ *(at least number restriction)*.

Note that $\top$ ($\bot$) can also be expressed as $\mathsf{C} \sqcup \neg \mathsf{C}$ ($\mathsf{C} \sqcap \neg \mathsf{C}$). For an arbitrary role R, the term $\exists_{\geq 1} \mathsf{R}$ can be rewritten as $\exists \mathsf{R} \,.\, \top$, $\exists_{\geq 0} \mathsf{R}$ as $\top$, and $\exists_{\leq 0} \mathsf{R}$ as $\forall \mathsf{R} \,.\, \bot$. Thus, we do not consider these terms as number restrictions in our language.

The concept language is syntactically restricting the combination of number restrictions and transitive roles. Number restrictions are only allowed for *simple* roles. This restriction is motivated by doubtful semantics for an unrestricted combinability and a simplified tableaux decision procedure. Moreover, this decision is supported by a recent undecidability result for the logic $\mathcal{ALCHNI}_{R^+}$ in case of an unrestricted combinability [Horrocks et al., 1999b].

Definition 4 (Generalized Concept Inclusions) If C and D are concept terms, then $\mathsf{C} \sqsubseteq \mathsf{D}$ *(generalized concept inclusion* or *GCI)* is a terminological axiom as well.

A finite set of terminological axioms $\mathcal{T}$ is called a *terminology* or *TBox*. GCIs can be used to represent terminological cycles. There exist at least two ways to deal with GCIs in a tableaux calculus. The 'internalization' approach (e.g. see in [Horrocks and Sattler, 1999]) makes use of the fact that the expressiveness of GCIs is already implied by the combination of role hierarchies and transitive roles. For instance, this allows one to introduce an internal transitive role U as a superrole of all other roles. Then, a GCI $\mathsf{C} \sqsubseteq \mathsf{D}$ can be internalized as $\forall \mathsf{U} . (\neg \mathsf{C} \sqcup \mathsf{D})$ and there is no need to adapt an tableaux calculus w.r.t. GCIs. However, with the presence of arbitrary ABoxes one has also to consider unrelated individuals. For instance, in this case the internalization approach could introduce a new internal root individual that is related with every other individual in the ABox via the superrole U. Then, the all-concepts corresponding to the internalized GCIs are added to the root individual. Alternatively, one could directly add the corresponding assertions to all ABox individuals (e.g. $\mathsf{i} : \forall \mathsf{U} . (\neg \mathsf{C} \sqcup \mathsf{D})$) instead of creating a root individual. We decided to pursue a different and more direct approach that extends an ABox tableaux calculus by new constructs and rules directly dealing with GCIs (see Definition 7).

The next definition gives a set-theoretic semantics to the language introduced above.

Definition 5 (Semantics) An *interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ consists of a set $\Delta^{\mathcal{I}}$ (the domain) and an interpretation function $\cdot^{\mathcal{I}}$. The interpretation function maps each concept name C to a subset $\mathsf{C}^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$, each role name R to a subset $\mathsf{R}^{\mathcal{I}}$ of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. Let the symbols C, D be concept expressions, and R, S be role names. Then the interpretation function can be extended to arbitrary concept and role terms as follows ($\| \cdot \|$ denotes the cardinality of a set):

$$
\begin{aligned}
(\mathsf{C} \sqcap \mathsf{D})^{\mathcal{I}} &:= \mathsf{C}^{\mathcal{I}} \cap \mathsf{D}^{\mathcal{I}} \\
(\mathsf{C} \sqcup \mathsf{D})^{\mathcal{I}} &:= \mathsf{C}^{\mathcal{I}} \cup \mathsf{D}^{\mathcal{I}} \\
(\neg \mathsf{C})^{\mathcal{I}} &:= \Delta^{\mathcal{I}} \setminus \mathsf{C}^{\mathcal{I}} \\
(\exists \mathsf{R} . \mathsf{C})^{\mathcal{I}} &:= \{a \in \Delta^{\mathcal{I}} \mid \exists b \in \Delta^{\mathcal{I}} : (a,b) \in \mathsf{R}^{\mathcal{I}}, b \in \mathsf{C}^{\mathcal{I}}\} \\
(\forall \mathsf{R} . \mathsf{C})^{\mathcal{I}} &:= \{a \in \Delta^{\mathcal{I}} \mid \forall b \in \Delta^{\mathcal{I}} : (a,b) \in \mathsf{R}^{\mathcal{I}} \Rightarrow b \in \mathsf{C}^{\mathcal{I}}\} \\
(\exists_{\geq n} \mathsf{R})^{\mathcal{I}} &:= \{a \in \Delta^{\mathcal{I}} \mid \; \|\{b \mid (a,b) \in \mathsf{R}^{\mathcal{I}}\}\| \; \geq n\} \\
(\exists_{\leq n} \mathsf{R})^{\mathcal{I}} &:= \{a \in \Delta^{\mathcal{I}} \mid \; \|\{b \mid (a,b) \in \mathsf{R}^{\mathcal{I}}\}\| \; \leq n\}
\end{aligned}
$$

An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ iff it satisfies (1) $\mathsf{C}^{\mathcal{I}} \subseteq \mathsf{D}^{\mathcal{I}}$ for all terminological axioms (GCIs) $\mathsf{C} \sqsubseteq \mathsf{D}$ in $\mathcal{T}$ and $\mathsf{R}^{\mathcal{I}} \subseteq \mathsf{S}^{\mathcal{I}}$ for all terminological axioms $\mathsf{R} \sqsubseteq \mathsf{S}$ (role inclusions) in $\mathcal{T}$, and (2) iff for every $\mathsf{R} \in T : \mathsf{R}^\mathcal{I} = (\mathsf{R}^\mathcal{I})^+$. A concept term C *subsumes* a concept term D w.r.t. a TBox $\mathcal{T}$ (written $\mathsf{D} \preceq_\mathcal{T} \mathsf{C}$), iff $\mathsf{D}^\mathcal{I} \subseteq \mathsf{C}^\mathcal{I}$ for all models $\mathcal{I}$ of $\mathcal{T}$. A concept term C is *satisfiable* w.r.t. a TBox $\mathcal{T}$ iff there exists a model $\mathcal{I}$ of $\mathcal{T}$ such that $\mathsf{C}^\mathcal{I} \neq \emptyset$.

One of the basic reasoning services for a description logic formalism is computing the subsumption relationship between named concepts (i.e. elements from C). This inference is needed in the TBox to build a hierarchy of concept names w.r.t. specificity. Satisfiability and subsumption can be mutually reduced to each other since $\mathsf{C} \preceq_\mathcal{T} \mathsf{D}$ iff $\mathsf{C} \sqcap \neg\mathsf{D}$ is not satisfiable w.r.t. $\mathcal{T}$ and C is unsatisfiable w.r.t. $\mathcal{T}$ iff $\mathsf{C} \preceq_\mathcal{T} \bot$.

2.3 The Assertional Language

In the following, the language for representing knowledge about individual worlds is introduced. An *ABox* $\mathcal{A}$ is a finite set of assertional axioms which are defined as follows:

Definition 6 (ABox Assertions) Let $O = O_O \cup O_N$ be a set of individual names, where the set O_O of "old" names is disjoint with the set O_N of "new" names. If C is a concept term, R a role name, and $\mathsf{a}, \mathsf{b} \in O$ are individual names, then the following expressions are *assertional axioms*:

- $\mathsf{a} : \mathsf{C}$ *(concept assertion)*,
- $(\mathsf{a}, \mathsf{b}) : \mathsf{R}$ *(role assertion)*.

The interpretation function $\cdot^\mathcal{I}$ of the interpretation $\mathcal{I}$ for the concept language can be extended to the assertional language by additionally mapping every individual name from O to a single element of $\Delta^\mathcal{I}$ in a way such that for $\mathsf{a}, \mathsf{b} \in O_O$, $\mathsf{a}^\mathcal{I} \neq \mathsf{b}^\mathcal{I}$ if $\mathsf{a} \neq \mathsf{b}$ *(unique name assumption)*. This ensures that different individuals in O_O are interpreted as different objects. The unique name assumption does not hold for elements of O_N, i.e. for $\mathsf{a}, \mathsf{b} \in O_N$, $\mathsf{a}^\mathcal{I} = \mathsf{b}^\mathcal{I}$ may hold even if $\mathsf{a} \neq \mathsf{b}$, or if we assume without loss of generality that $\mathsf{a} \in O_N, \mathsf{b} \in O_O$. An interpretation satisfies an assertional axiom $\mathsf{a} : \mathsf{C}$ iff $\mathsf{a}^\mathcal{I} \in \mathsf{C}^\mathcal{I}$ and $(\mathsf{a}, \mathsf{b}) : \mathsf{R}$ iff $(\mathsf{a}^\mathcal{I}, \mathsf{b}^\mathcal{I}) \in \mathsf{R}^\mathcal{I}$.

An interpretation is a *model* of an ABox $\mathcal{A}$ w.r.t. a TBox $\mathcal{T}$ iff it is a model of $\mathcal{T}$ and furthermore satisfies all assertional axioms in $\mathcal{A}$. An ABox is *consistent* w.r.t. a TBox $\mathcal{T}$ iff it has a model w.r.t. $\mathcal{T}$. An individual b is called a *direct successor* of an individual a in an ABox $\mathcal{A}$ iff $\mathcal{A}$ contains the assertional axiom $(\mathsf{a}, \mathsf{b}) : \mathsf{R}$. An individual b is called a *successor* of a if it is either a direct successor of a or there exists in $\mathcal{A}$ a chain of assertions $(\mathsf{a}, \mathsf{b_1}) : \mathsf{R_1}, (\mathsf{b_1}, \mathsf{b_2}) : \mathsf{R_2}, \ldots, (\mathsf{b_n}, \mathsf{b}) : \mathsf{R_{n-1}}$. In case that $\mathsf{R_i} = \mathsf{R_j}$ or $\mathsf{R_i} \in \mathsf{R}^\downarrow$ for all $\mathsf{i}, \mathsf{j} \in 1..\mathsf{n}-1$ we call b the (direct) R-*successor* of a. A (direct) *predecessor* is defined analogously. An individual a is called an *instance* of a concept term C in an interpretation $\mathcal{I}$ iff $\mathsf{a}^{\mathcal{I}} \in \mathsf{C}^{\mathcal{I}}$. The *direct types* of an individual are the most specific atomic concepts which the individual is an instance of.

The ABox consistency problem is to decide whether a given ABox $\mathcal{A}$ is consistent w.r.t. a TBox $\mathcal{T}$. Satisfiability of concept terms can be reduced to ABox consistency as follows: A concept term C is satisfiable iff the ABox $\{\mathsf{a} : \mathsf{C}\}$ is consistent. *Instance checking* tests whether an individual a is an instance of a concept term C w.r.t. an ABox $\mathcal{A}$ and a TBox $\mathcal{T}$, i.e. whether $\mathcal{A}$ entails $\mathsf{a} : \mathsf{C}$ w.r.t. $\mathcal{T}$. This problem is reduced to the problem of deciding if the ABox $\mathcal{A} \cup \{\mathsf{a} : \neg \mathsf{C}\}$ is inconsistent.

3 An ABox Example

Before we continue with the calculus for $\mathcal{ALCNH}_{R^+}$, we illustrate in the following the expressiveness of $\mathcal{ALCNH}_{R^+}$ with a TBox and ABox example about family relationships. This example uses prominent features of $\mathcal{ALCNH}_{R^+}$ such as transitive roles, role hierarchies, number restrictions and generalized concept inclusions.

In the TBox *family* we assume a role has_descendant which is declared to be *transitive*, has_gender which is declared as a feature (e.g. this can be achieved by adding the axiom $\top \sqsubseteq \exists_{\leq 1}$ has_gender), and a role has_sibling. The TBox *family* contains the following role axioms.

has_child $\sqsubseteq$ has_descendant

has_sister $\sqsubseteq$ has_sibling

has_brother $\sqsubseteq$ has_sibling

The TBox *family* contains concept axioms specifying the domain and/or range of the roles introduced above (the domain A of a role R can be expressed by the axiom $\exists_{\geq 1} \mathsf{R} \sqsubseteq \mathsf{A}$ and the range B by $\top \sqsubseteq \forall \mathsf{R} . \mathsf{B}$).

$\exists_{\geq 1}$ has_descendant $\sqsubseteq$ human

$\top \sqsubseteq \forall$ has_descendant . human

$\exists_{\geq 1}$ has_child $\sqsubseteq$ parent

$\exists_{\geq 1}$ has_sibling $\sqsubseteq$ sibling

$\top \sqsubseteq \forall$ has_sibling . sibling

$\top \sqsubseteq \forall$ has_sister . sister

$\top \sqsubseteq \forall$ has_brother . brother

$\top \sqsubseteq \forall$ has_gender . (female $\sqcup$ male)

The next axioms guarantee the disjointness between the concepts female, male, and human.

$\mathsf{female} \sqsubseteq \neg(\mathsf{human} \sqcup \mathsf{male})$

$\mathsf{male} \sqsubseteq \neg(\mathsf{human} \sqcup \mathsf{female})$

$\mathsf{human} \sqsubseteq \neg(\mathsf{female} \sqcup \mathsf{male})$

After these preliminaries we start with axioms expressing basic knowledge about family members. We use $\mathsf{C} \doteq \mathsf{D}$ as an abbreviation for $\mathsf{C} \sqsubseteq \mathsf{D}$ and $\mathsf{D} \sqsubseteq \mathsf{C}$.

$\mathsf{human} \sqsubseteq \exists_{\geq 1}\, \mathsf{has\_gender}$

$\mathsf{woman} \doteq \mathsf{human} \sqcap \forall\, \mathsf{has\_gender} . \mathsf{female}$

$\mathsf{man} \doteq \mathsf{human} \sqcap \forall\, \mathsf{has\_gender} . \mathsf{male}$

$\mathsf{parent} \doteq \exists_{\geq 1}\, \mathsf{has\_child}$

$\mathsf{mother} \doteq \mathsf{woman} \sqcap \mathsf{parent}$

$\mathsf{father} \doteq \mathsf{man} \sqcap \mathsf{parent}$

The next axioms describe some aspects of relatives of a family. Note the inferred equivalences between the concept pairs "mother_with_..." and "mother_having_..." as shown in Figure 1.

$\mathsf{mother\_having\_only\_female\_kids} \doteq \mathsf{mother} \sqcap \forall\, \mathsf{has\_child} . \forall\, \mathsf{has\_gender} . \mathsf{female}$

$\mathsf{mother\_having\_only\_daughters} \doteq \mathsf{mother} \sqcap \exists_{\geq 1}\, \mathsf{has\_child} \sqcap \forall\, \mathsf{has\_child} . \mathsf{woman}$

$\mathsf{mother\_with\_kids} \doteq \mathsf{mother} \sqcap \exists_{\geq 2}\, \mathsf{has\_child}$

$\mathsf{grandpa} \doteq \mathsf{man} \sqcap \exists\, \mathsf{has\_child} . \mathsf{parent}$

$\mathsf{great\_grandpa} \doteq \mathsf{man} \sqcap \exists\, \mathsf{has\_child} . (\exists\, \mathsf{has\_child} . \mathsf{parent})$

$\mathsf{grandma} \doteq \mathsf{woman} \sqcap \exists\, \mathsf{has\_child} . \mathsf{parent}$

$\mathsf{great\_grandma} \doteq \mathsf{woman} \sqcap \exists\, \mathsf{has\_child} . (\exists\, \mathsf{has\_child} . \mathsf{parent})$

$\mathsf{aunt} \doteq \mathsf{woman} \sqcap \exists\, \mathsf{has\_sibling} . \mathsf{parent}$

$\mathsf{uncle} \doteq \mathsf{man} \sqcap \exists\, \mathsf{has\_sibling} . \mathsf{parent}$

$\mathsf{sibling} \doteq \mathsf{sister} \sqcup \mathsf{brother}$

$\mathsf{sister} \doteq \mathsf{woman} \sqcap \exists_{\geq 1}\, \mathsf{has\_sibling}$

$\mathsf{brother} \doteq \mathsf{man} \sqcap \exists_{\geq 1}\, \mathsf{has\_sibling}$

$\mathsf{mother\_with\_siblings} \doteq \mathsf{mother} \sqcap \forall\, \mathsf{has\_child} . \mathsf{sibling}$

There still exists no formal relationship between the notions "having kids" and "having siblings." This is expressed by the next two axioms. The last axiom defines a concept mother_having_only_sisters which has the other specific

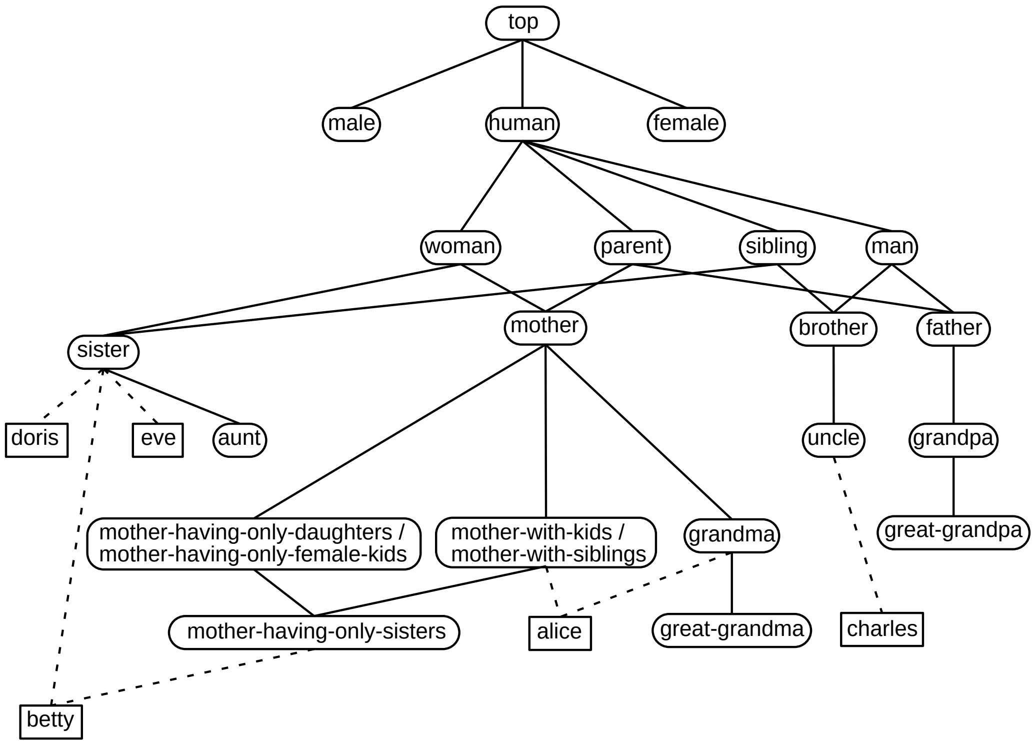

Figure 1: Concept hierarchy of the TBox *family* augmented with the individuals from the ABox *smith_family*. Ovals represent atomic concepts, rectangles denote ABox individuals, solid lines show the direct subsumption relationship, and dashed lines the instance membership of the individuals for their direct types.

"mother_..." concepts as parents (see Figure 1).

$$\exists_{\geq 2}\ \mathsf{has\_child} \sqsubseteq\ \forall\, \mathsf{has\_child} \,.\, \mathsf{sibling}$$

$$\exists\, \mathsf{has\_child} \,.\, \mathsf{sibling} \sqsubseteq\ \exists_{\geq 2}\ \mathsf{has\_child}$$

$$\mathsf{mother\_having\_only\_sisters} \doteq\ \mathsf{mother} \sqcap \forall\, \mathsf{has\_child} \,.\, (\mathsf{sister} \sqcap \forall\, \mathsf{has\_sibling} \,.\, \mathsf{sister})$$

Using the TBox *family*, the ABox *smith_family* is specified. It consists of several assertions about the individuals alice, betty, charles, doris, and eve. The individual alice is the mother of her two children betty and charles.

$$\begin{aligned} \mathsf{alice} &: \mathsf{woman} \sqcap \exists_{\leq 2}\ \mathsf{has\_child} \\ (\mathsf{alice}, \mathsf{betty}) &: \mathsf{has\_child} \\ (\mathsf{alice}, \mathsf{charles}) &: \mathsf{has\_child} \end{aligned}$$

The individual betty is the sibling of charles and the mother of doris and eve, who are the only siblings of each other. The individual charles is the only brother of betty.

$$
\begin{aligned}
\mathsf{betty} &: \mathsf{woman} \sqcap \exists_{\leq 2}\, \mathsf{has\_child} \sqcap \exists_{\leq 1}\, \mathsf{has\_sibling} \\
(\mathsf{betty}, \mathsf{doris}) &: \mathsf{has\_child} \\
(\mathsf{betty}, \mathsf{eve}) &: \mathsf{has\_child} \\
(\mathsf{betty}, \mathsf{charles}) &: \mathsf{has\_sibling} \\
\mathsf{charles} &: \mathsf{brother} \sqcap \exists_{\leq 1}\, \mathsf{has\_sibling} \\
(\mathsf{charles}, \mathsf{betty}) &: \mathsf{has\_sibling} \\
\mathsf{doris} &: \exists_{\leq 1}\, \mathsf{has\_sibling} \\
\mathsf{eve} &: \exists_{\leq 1}\, \mathsf{has\_sibling} \\
(\mathsf{doris}, \mathsf{eve}) &: \mathsf{has\_sister} \\
(\mathsf{eve}, \mathsf{doris}) &: \mathsf{has\_sister}
\end{aligned}
$$

Figure 1 also shows the inferred *direct types* of the individuals in ABox *smith_family*. alice has as direct types {mother_with_siblings, grandma}, betty has {mother_having_only_sisters, sister}, charles has {uncle}, and doris and eve have {sister}. These inferences demonstrate the expressiveness of $\mathcal{ALCNH}_{R^+}$. The ABox *smith_family* contains only minimal knowledge about the individuals and their relationships.

4 A Tableaux Calculus for $\mathcal{ALCNH}_{R^+}$

In the following we devise a *tableaux* algorithm to decide the consistency of $\mathcal{ALCNH}_{R^+}$ ABoxes. The algorithm is characterized by a set of tableaux or *completion* rules and by a particular *completion strategy* ensuring a specific order for applying the completion rules to assertional axioms of an ABox. The strategy is essential to guarantee the completeness of the ABox consistency algorithm. First, we have to introduce new assertional axioms needed to define the augmentation of an ABox.

Definition 7 (Additional ABox Assertions) Let C be a concept term, the individual names $\mathsf{a}, \mathsf{b} \in O$, and $x \notin O$, then the following expressions are also assertional axioms:

- $\forall\, x \,.\, x : \mathsf{C}$ *(universal concept assertion)*,[1]
- $\mathsf{a} \neq \mathsf{b}$ *(inequality assertion)*.

An interpretation $\mathcal{I}$ satisfies an assertional axiom $\forall\, x \,.\, x : \mathsf{C}$ iff $\mathsf{C}^{\mathcal{I}} = \Delta^{\mathcal{I}}$ and $\mathsf{a} \neq \mathsf{b}$ iff $\mathsf{a}^{\mathcal{I}} \neq \mathsf{b}^{\mathcal{I}}$.

[1] $\forall\, x \,.\, x : \mathsf{C}$ should be read as $\forall\, x \,.\, (x : \mathsf{C})$.

Given the new ABox assertions we define for any concept term its negation normal form.

Definition 8 (Negation Normal Form) We assume the same naming conventions as in Definition 3. The negation normal form is defined by applying the following transformations in such a way that a negation sign may occur only in front of concept names. This transformation is possible in linear time.

- $\neg\top \equiv \bot$
- $\neg\bot \equiv \top$
- $\neg(\mathsf{C} \sqcap \mathsf{D}) \equiv \neg\mathsf{C} \sqcup \neg\mathsf{D}$
- $\neg(\mathsf{C} \sqcup \mathsf{D}) \equiv \neg\mathsf{C} \sqcap \neg\mathsf{D}$
- $\neg\forall\,\mathsf{R}\,.\,\mathsf{C} \equiv \exists\,\mathsf{R}\,.\,\neg\mathsf{C}$
- $\neg\exists\,\mathsf{R}\,.\,\mathsf{C} \equiv \forall\,\mathsf{R}\,.\,\neg\mathsf{C}$
- $\neg\exists_{\leq m}\,\mathsf{S} \equiv \exists_{\geq m+1}\,\mathsf{S}$
- $\neg\exists_{\geq m}\,\mathsf{S} \equiv \exists_{\leq m-1}\,\mathsf{S}$

We are now ready to define an augmented ABox as input to the tableaux rules.

Definition 9 (Augmented ABox) For an initial ABox $\mathcal{A}$ w.r.t a TBox $\mathcal{T}$ we define its *augmented* ABox $\mathcal{A}'$ by applying the following rules to $\mathcal{A}$. For every GCI $\mathsf{C} \sqsubseteq \mathsf{D}$ in $\mathcal{T}$ the assertion $\forall\,x\,.\,x\!:\!(\neg\mathsf{C} \sqcup \mathsf{D})$ is added to $\mathcal{A}'$. Every concept term occurring in $\mathcal{A}$ is transformed into its negation normal form. Let $O_O := \{\mathsf{a}_1, \ldots, \mathsf{a}_\mathsf{n}\}$ be the set of old individual names mentioned in $\mathcal{A}$, then the set of inequality assertions $\{\mathsf{a}_\mathsf{i} \not\doteq \mathsf{a}_\mathsf{j} \mid \mathsf{a}_\mathsf{i}, \mathsf{a}_\mathsf{j} \in O_O, \mathsf{i},\mathsf{j} \in 1..\mathsf{n}, \mathsf{i} \neq \mathsf{j}\}$ is added to $\mathcal{A}$. From this point on, if we refer to an initial ABox $\mathcal{A}$ we always mean its augmented ABox.

The tableaux rules also require the notion of *blocking* their applicability. This is based on so-called concept sets.

Definition 10 (Concept Set, $\mathcal{A}$-equivalent, $\mathcal{A}$-blocked) Given an ABox $\mathcal{A}$ and an individual a occurring in $\mathcal{A}$, we define the *concept set* of a as $\sigma(\mathcal{A}, \mathsf{a}) := \{\top\} \cup \{\mathsf{C} \mid \mathsf{a}\!:\!\mathsf{C} \in \mathcal{A}\}$. We define two individuals as *$\mathcal{A}$-equivalent*, written $\mathsf{a} \equiv_\mathcal{A} \mathsf{b}$, if their concept sets are equal, i.e. $\sigma(\mathcal{A}, \mathsf{a}) = \sigma(\mathcal{A}, \mathsf{b})$. We say that an individual b is *$\mathcal{A}$-blocked*[2] by a, written $\mathsf{a} \succcurlyeq_\mathcal{A} \mathsf{b}$, if $\sigma(\mathcal{A}, \mathsf{a}) \supseteq \sigma(\mathcal{A}, \mathsf{b})$.

4.1 Completion Rules

We are now ready to define the *completion rules* that are intended to generate a so-called completion of an ABox (see also below).

[2]We may omit the reference to $\mathcal{A}$ by speaking of *blocked* if the context is obvious.

Definition 11 (Completion Rules)

R⊓ The conjunction rule.
if 1. $\mathsf{a}:\mathsf{C}\sqcap\mathsf{D}\in\mathcal{A}$, and
2. $\{\mathsf{a}:\mathsf{C},\ \mathsf{a}:\mathsf{D}\}\not\subseteq\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{a}:\mathsf{C},\ \mathsf{a}:\mathsf{D}\}$

R⊔ The disjunction rule (nondeterministic).
if 1. $\mathsf{a}:\mathsf{C}\sqcup\mathsf{D}\in\mathcal{A}$, and
2. $\{\mathsf{a}:\mathsf{C},\ \mathsf{a}:\mathsf{D}\}\cap\mathcal{A}=\emptyset$
then $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{a}:\mathsf{C}\}$ **or** $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{a}:\mathsf{D}\}$

R∀C The role value restriction rule.
if 1. $\mathsf{a}:\forall\,\mathsf{R}\,.\,\mathsf{C}\in\mathcal{A}$, and
2. $\exists\,\mathsf{b}\in O,\mathsf{S}\in\mathsf{R}^{\downarrow}:(\mathsf{a},\mathsf{b}):\mathsf{S}\in\mathcal{A}$, and
3. $\mathsf{b}:\mathsf{C}\notin\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{b}:\mathsf{C}\}$

R$\forall_+$C The transitive role value restriction rule.
if 1. $\mathsf{a}:\forall\,\mathsf{R}\,.\,\mathsf{C}\in\mathcal{A}$, and
2. $\exists\,\mathsf{b}\in O,\mathsf{T}\in\mathsf{R}^{\downarrow},\mathsf{T}\in T,\mathsf{S}\in\mathsf{T}^{\downarrow}:(\mathsf{a},\mathsf{b}):\mathsf{S}\in\mathcal{A}$, and
3. $\mathsf{b}:\forall\,\mathsf{T}\,.\,\mathsf{C}\notin\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{b}:\forall\,\mathsf{T}\,.\,\mathsf{C}\}$

R$\forall_x$ The universal concept restriction rule.
if 1. $\forall\,x\,.\,x:\mathsf{C}\in\mathcal{A}$, and
2. $\exists\,\mathsf{a}\in O$: a mentioned in $\mathcal{A}$, and
3. $\mathsf{a}:\mathsf{C}\notin\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{\mathsf{a}:\mathsf{C}\}$

R∃C The role exists restriction rule (generating).
if 1. $\mathsf{a}:\exists\,\mathsf{R}\,.\,\mathsf{C}\in\mathcal{A}$, and
2. $\mathsf{a}\in O_N\Rightarrow(\neg\exists\,\mathsf{c}\in O_N:\mathsf{c}$ mentioned in $\mathcal{A},\mathsf{c}\not\succcurlyeq_{\mathcal{A}}\mathsf{a})$, and
3. $\neg\exists\,\mathsf{b}\in O,\mathsf{S}\in\mathsf{R}^{\downarrow}:\{(\mathsf{a},\mathsf{b}):\mathsf{S},\ \mathsf{b}:\mathsf{C}\}\subseteq\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{(\mathsf{a},\mathsf{b}):\mathsf{R},\ \mathsf{b}:\mathsf{C}\}$ where $\mathsf{b}\in O_N$ is not used in $\mathcal{A}$

R$\exists_{\geq n}$ The number restriction exists rule (generating).
if 1. $\mathsf{a}:\exists_{\geq n}\,\mathsf{R}\in\mathcal{A}$, and
2. $\mathsf{a}\in O_N\Rightarrow(\neg\exists\,\mathsf{c}\in O_N:\mathsf{c}$ mentioned in $\mathcal{A},\mathsf{c}\not\succcurlyeq_{\mathcal{A}}\mathsf{a})$, and
3. $\neg\exists\,\mathsf{b}_1,\ldots,\mathsf{b}_\mathsf{n}\in O,\mathsf{S}_1,\ldots,\mathsf{S}_\mathsf{n}\in\mathsf{R}^{\downarrow}:$
$\{(\mathsf{a},\mathsf{b}_\mathsf{k}):\mathsf{S}_\mathsf{k}\,|\,\mathsf{k}\in 1..\mathsf{n}\}\cup\{\mathsf{b}_\mathsf{i}\not\doteq\mathsf{b}_\mathsf{j}\,|\,\mathsf{i},\mathsf{j}\in 1..\mathsf{n},\mathsf{i}\neq\mathsf{j}\}\subseteq\mathcal{A}$
then $\mathcal{A}'=\mathcal{A}\cup\{(\mathsf{a},\mathsf{b}_\mathsf{k}):\mathsf{R}\,|\,\mathsf{k}\in 1..\mathsf{n}\}\cup\{\mathsf{b}_\mathsf{i}\not\doteq\mathsf{b}_\mathsf{j}\,|\,\mathsf{i},\mathsf{j}\in 1..\mathsf{n},\mathsf{i}\neq\mathsf{j}\}$
where $\mathsf{b}_1,\ldots,\mathsf{b}_\mathsf{n}\in O_N$ are not used in $\mathcal{A}$

R$\exists_{\leq n}$ The number restriction merge rule (nondeterministic).

if 1. $\mathsf{a} : \exists_{\leq n} \mathsf{R} \in \mathcal{A}$, and

2. $\exists\, \mathsf{b_1}, \dots, \mathsf{b_m} \in O, \mathsf{S_1}, \dots, \mathsf{S_m} \in \mathsf{R}^{\downarrow}\colon \{(\mathsf{a}, \mathsf{b_1}) : \mathsf{S_1}, \dots, (\mathsf{a}, \mathsf{b_m}) : \mathsf{S_m}\} \subseteq \mathcal{A}$ with $\mathsf{m} > \mathsf{n}$, and

3. $\exists\, \mathsf{b_i}, \mathsf{b_j} \in \{\mathsf{b_1}, \dots, \mathsf{b_m}\} : \mathsf{i} \neq \mathsf{j},\ \mathsf{b_i} \neq \mathsf{b_j} \notin \mathcal{A}$

then $\mathcal{A}' = \mathcal{A}[\mathsf{b_i}/\mathsf{b_j}]$, i.e. replace every occurrence of $\mathsf{b_i}$ in $\mathcal{A}$ by $\mathsf{b_j}$

We call the rules R$\sqcup$ and R$\exists_{\leq n}$ *nondeterministic* rules since they can be applied in different ways to the same ABox. The remaining rules are called *deterministic* rules. Moreover, we call the rules R$\exists$C and R$\exists_{\geq n}$ *generating* rules since they are the only rules that introduce new individuals in an ABox.

The increase of expressiveness in $\mathcal{ALCNH}_{R^+}$ gained by supporting ABox reasoning is reflected in tableaux rules that are more complex than in comparable approaches for concept consistency. The universal concept restriction rule takes care of GCIs and usually causes additional complexity by adding disjunctions to an ABox. The generating rules have a more complex premise since they may test only for a blocking situation if they are applied to new individuals, i.e. a blocking situation can never occur for old individuals. The necessity of this additional precondition is illustrated by the following example. We define a concept D where R is a transitive superrole of S.

$$\mathsf{D} \doteq \mathsf{C} \sqcap \exists \mathsf{S} . \mathsf{C} \sqcap \exists_{\leq 1} \mathsf{S} \sqcap \forall \mathsf{R} . \exists \mathsf{S} . \mathsf{C}$$

$$\mathcal{A} := \{(\mathsf{i}, \mathsf{j}) : \mathsf{S},\ (\mathsf{j}, \mathsf{k}) : \mathsf{S},\ \mathsf{i} : \mathsf{D},\ \mathsf{j} : \mathsf{D},\ \mathsf{k} : \neg \mathsf{C}\}$$

Then, we define an ABox $\mathcal{A}$ which is obviously unsatisfiable due to a clash for the individual k with $\mathsf{C} \sqcap \neg\mathsf{C}$. However, if blocking were allowed for old individuals, the role exists restriction rule would not create a S-successor with qualification C for the individual j. As a consequence, the number restriction merge rule would never merge this successor with the individual k which results in the unsatisfiability of $\mathcal{A}$.

Proposition 12 (Invariance) Let $\mathcal{A}$ and $\mathcal{A}'$ be ABoxes. Then:

1. If $\mathcal{A}'$ is derived from $\mathcal{A}$ by applying a deterministic rule, then $\mathcal{A}$ is satisfiable iff $\mathcal{A}'$ is satisfiable.

2. If $\mathcal{A}'$ is derived from $\mathcal{A}$ by applying a nondeterministic rule, then $\mathcal{A}$ is satisfiable if $\mathcal{A}'$ is satisfiable. Conversely, if $\mathcal{A}$ is satisfiable and a nondeterministic rule is applicable to $\mathcal{A}$, then it can be applied in such a way that it yields a satisfiable ABox $\mathcal{A}'$.

Proof. **1.** "$\Leftarrow$" Due to the structure of the deterministic rules one can immediately verify that $\mathcal{A}$ is a subset of $\mathcal{A}'$. Therefore, $\mathcal{A}$ is satisfiable if $\mathcal{A}'$ is satisfiable.

"$\Rightarrow$" In order to show that $\mathcal{A}'$ is satisfiable after applying a deterministic rule to the satisfiable ABox $\mathcal{A}$, we examine each applicable rule separately. We assume that $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ satisfies $\mathcal{A}$.

If the conjunction rule is applied to $\mathsf{a} : \mathsf{C} \sqcap \mathsf{D} \in \mathcal{A}$, then we get a new Abox $\mathcal{A}' = \mathcal{A} \cup \{\mathsf{a} : \mathsf{C},\ \mathsf{a} : \mathsf{D}\}$. Since $\mathcal{I}$ satisfies $\mathsf{a} : \mathsf{C} \sqcap \mathsf{D}$, $\mathcal{I}$ satisfies $\mathsf{a} : \mathsf{C}$ and $\mathsf{a} : \mathsf{D}$ and therefore $\mathcal{A}'$.

If the role value restriction rule is applied to $\mathsf{a} : \forall\, \mathsf{R} \,.\, \mathsf{C} \in \mathcal{A}$, then there must be a role assertion $(\mathsf{a}, \mathsf{b}) : \mathsf{S} \in \mathcal{A}$ with $\mathsf{S} \in \mathsf{R}^{\downarrow}$ such that $\mathcal{A}' = \mathcal{A} \cup \{\mathsf{b} : \mathsf{C}\}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, it holds that $(\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}) \in \mathsf{S}^{\mathcal{I}} \subseteq \mathsf{R}^{\mathcal{I}}$. Since $\mathcal{I}$ satisfies $\mathsf{a} : \forall\, \mathsf{R} \,.\, \mathsf{C}$, it holds that $\mathsf{b}^{\mathcal{I}} \in \mathsf{C}^{\mathcal{I}}$. Thus, $\mathcal{I}$ satisfies $\mathsf{b} : \mathsf{C}$ and therefore $\mathcal{A}'$.

If the transitive role value restriction rule is applied to $\mathsf{a} : \forall\, \mathsf{R} \,.\, \mathsf{C} \in \mathcal{A}$, there must be an assertion $(\mathsf{a}, \mathsf{b}) : \mathsf{S} \in \mathcal{A}$ with $\mathsf{S} \in \mathsf{T}^{\downarrow} \subseteq \mathsf{R}^{\downarrow}$, $\mathsf{T} \in \mathcal{T}$ such that we get $\mathcal{A}' = \mathcal{A} \cup \{\mathsf{b} : \forall\, \mathsf{T} \,.\, \mathsf{C}\}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, we have $\mathsf{a}^{\mathcal{I}} \in (\forall\, \mathsf{R} \,.\, \mathsf{C})^{\mathcal{I}}$ and $(\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}) \in \mathsf{S}^{\mathcal{I}} \subseteq \mathsf{T}^{\mathcal{I}} \subseteq \mathsf{R}^{\mathcal{I}}$. Since $\mathcal{I}$ satisfies $\mathsf{a} : \forall\, \mathsf{T} \,.\, \mathsf{C}$ and $\mathsf{T} \in \mathcal{T}, \mathsf{T} \in \mathsf{R}^{\downarrow}$, it holds that $\mathsf{b}^{\mathcal{I}} \in (\forall\, \mathsf{T} \,.\, \mathsf{C})^{\mathcal{I}}$ unless there exists a successor c of b such that $(\mathsf{b}, \mathsf{c}) : \mathsf{S}' \in \mathcal{A}$, $(\mathsf{b}^{\mathcal{I}}, \mathsf{c}^{\mathcal{I}}) \in \mathsf{S}'^{\mathcal{I}} \subseteq \mathsf{T}^{\mathcal{I}}$ and $\mathsf{c}^{\mathcal{I}} \notin \mathsf{C}^{\mathcal{I}}$. It follows from $(\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}) \in \mathsf{T}^{\mathcal{I}}$, $(\mathsf{b}^{\mathcal{I}}, \mathsf{c}^{\mathcal{I}}) \in \mathsf{T}^{\mathcal{I}}$, and $\mathsf{T} \in \mathcal{T}$ that $(\mathsf{a}^{\mathcal{I}}, \mathsf{c}^{\mathcal{I}}) \in \mathsf{T}^{\mathcal{I}} \subseteq \mathsf{R}^{\mathcal{I}}$ and $\mathsf{a}^{\mathcal{I}} \notin (\forall\, \mathsf{R} \,.\, \mathsf{C})^{\mathcal{I}}$ in contradiction to the assumption. Thus, $\mathcal{I}$ satisfies $\mathsf{b} : \forall\, \mathsf{T} \,.\, \mathsf{C}$ and therefore $\mathcal{A}'$.

If the universal concept restriction rule is applied to an individual a in $\mathcal{A}$ because of $\forall\, x \,.\, x : \mathsf{C} \in \mathcal{A}$, then $\mathcal{A}' = \mathcal{A} \cup \{\mathsf{a} : \mathsf{C}\}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, it holds that $\mathsf{C}^{\mathcal{I}} = \Delta^{\mathcal{I}}$. Thus, it holds that $\mathsf{a}^{\mathcal{I}} \in \mathsf{C}^{\mathcal{I}}$ and $\mathcal{I}$ satisfies $\mathcal{A}'$.

If the role exists restriction rule is applied to $\mathsf{a} : \exists\, \mathsf{R} \,.\, \mathsf{C} \in \mathcal{A}$, then we get the ABox $\mathcal{A}' = \mathcal{A} \cup \{(\mathsf{a}, \mathsf{b}) : \mathsf{R},\ \mathsf{b} : \mathsf{C}\}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, there exists a $y \in \Delta^{\mathcal{I}}$ such that $(\mathsf{a}^{\mathcal{I}}, y) \in \mathsf{R}^{\mathcal{I}}$ and $y \in \mathsf{C}^{\mathcal{I}}$. We define the interpretation function $\cdot^{\mathcal{I}'}$ such that $\mathsf{b}^{\mathcal{I}'} := y$ and $x^{\mathcal{I}'} := x^{\mathcal{I}}$ for $x \neq \mathsf{b}$. It is easy to show that $\mathcal{I}' = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}'})$ satisfies $\mathcal{A}'$.

If the number restriction exists rule is applied to $\mathsf{a} : \exists_{\geq n}\, \mathsf{R} \in \mathcal{A}$, then we get $\mathcal{A}' = \mathcal{A} \cup \{(\mathsf{a}, \mathsf{b_k}) : \mathsf{R} \,|\, \mathsf{k} \in 1..\mathsf{n}\} \cup \{\mathsf{b_i} \neq \mathsf{b_j} \,|\, \mathsf{i}, \mathsf{j} \in 1..\mathsf{n}, \mathsf{i} \neq \mathsf{j}\}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, there must exist n distinct individuals $y_i \in \Delta^{\mathcal{I}}$, $i \in 1..n$ such that $(\mathsf{a}^{\mathcal{I}}, y_i) \in \mathsf{R}^{\mathcal{I}}$. We define the interpretation function $\cdot^{\mathcal{I}'}$ such that $\mathsf{b_i}^{\mathcal{I}'} := y_i$ and $x^{\mathcal{I}'} := x^{\mathcal{I}}$ for $x \notin \{\mathsf{b_1}, \ldots, \mathsf{b_n}\}$. It is easy to show that $\mathcal{I}' = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}'})$ satisfies $\mathcal{A}'$.

2. "$\Leftarrow$" Assume that $\mathcal{A}'$ is satisfied by $\mathcal{I}' = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}'})$. We show that $\mathcal{A}$ is also satisfiable by examining the nondeterministic rules.

If $\mathcal{A}'$ is obtained from $\mathcal{A}$ by applying the disjunction rule, then $\mathcal{A}$ is a subset of $\mathcal{A}'$ and therefore satisfied by $\mathcal{I}'$.

If $\mathcal{A}'$ is obtained from $\mathcal{A}$ by applying the number restriction merge rule to $\mathsf{a} : \exists_{\leq n}\, \mathsf{R} \in \mathcal{A}$, then there exist $\mathsf{b_i}, \mathsf{b_j}$ in $\mathcal{A}$ such that $\mathcal{A}' = \mathcal{A}[\mathsf{b_i}/\mathsf{b_j}]$. We define the interpretation function $\cdot^{\mathcal{I}}$ such that $\mathsf{b_i}^{\mathcal{I}} := \mathsf{b_j}^{\mathcal{I}'}$ and $\mathsf{x}^{\mathcal{I}} := \mathsf{x}^{\mathcal{I}'}$ for every $\mathsf{x} \neq \mathsf{b_i}$. Obviously $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ satisfies $\mathcal{A}$.

"⇒" We suppose that $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ satisfies $\mathcal{A}$ and a nondeterministic rule is applicable to an individual a in $\mathcal{A}$.

If the disjunction rule is applicable to $\mathsf{a} : \mathsf{C} \sqcup \mathsf{D} \in \mathcal{A}$ and $\mathcal{A}$ is satisfiable, it holds $\mathsf{a}^{\mathcal{I}} \in (\mathsf{C} \sqcup \mathsf{D})^{\mathcal{I}}$. It follows that either $\mathsf{a}^{\mathcal{I}} \in \mathsf{C}^{\mathcal{I}}$ or $\mathsf{a}^{\mathcal{I}} \in \mathsf{D}^{\mathcal{I}}$ (or both). Hence, the disjunction rule can be applied in a way that $\mathcal{I}$ also satisfies the ABox $\mathcal{A}'$.

If the number restriction merge rule is applicable to $\mathsf{a} : \exists_{\leq n} \mathsf{R} \in \mathcal{A}$ and $\mathcal{A}$ is satisfiable, it holds $\mathsf{a}^{\mathcal{I}} \in (\exists_{\leq n} \mathsf{R})^{\mathcal{I}}$ and $\|\{b \mid (a, b) \in \mathsf{R}^{\mathcal{I}}\}\| \leq n$. However, it also holds $\|\{\mathsf{b} \mid (\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}) \in \mathsf{R}^{\mathcal{I}}\}\| > m$ with $m \geq n$.[3] Thus, we can conclude by the Pigeonhole Principle (e.g. see [Lewis and Papadimitriou, 1981, page 26]) that there exist at least two R-successors $\mathsf{b_i}, \mathsf{b_j}$ of a such that $\mathsf{b_i}^{\mathcal{I}} = \mathsf{b_j}^{\mathcal{I}}$. Since $\mathcal{I}$ satisfies $\mathcal{A}$, we have $\mathsf{b_i} \neq \mathsf{b_j} \notin \mathcal{A}$ and at least one of the two individuals must be a new individual. Let us assume that $\mathsf{b_i} \in O_N$ and $\mathsf{b_i} = \mathsf{b_j}$, then $\mathcal{I}$ obviously satisfies $\mathcal{A}[\mathsf{b_i}/\mathsf{b_j}]$. □

Given an initial ABox $\mathcal{A}$, more than one rule might be applicable to $\mathcal{A}$. This is controlled by a completion strategy in accordance to an ordering for new individuals.

Definition 13 (Individual Ordering) We define an *individual ordering* '$\prec$' for new individuals (elements of O_N) occurring in an ABox $\mathcal{A}$. If $\mathsf{b} \in O_N$ is introduced in $\mathcal{A}$, then $\mathsf{a} \prec \mathsf{b}$ for all new individuals a already present in $\mathcal{A}$.

Definition 14 (Completion Strategy) We define a *completion strategy* that must observe the following restrictions.

- Meta rules:
  - Apply a rule to an individual $\mathsf{b} \in O_N$ only if no rule is applicable to an individual $\mathsf{a} \in O_O$.
  - Apply a rule to an individual $\mathsf{b} \in O_N$ only if no rule is applicable to another individual $\mathsf{a} \in O_N$ such that $\mathsf{a} \prec \mathsf{b}$.
- The completion rules are always applied in the following order. A step is skipped in case the corresponding set of applicable rules is empty.
  1. Apply all nongenerating rules (R$\sqcap$, R$\sqcup$, R$\forall$C, R$\forall_+$C, R$\forall_x$, R$\exists_{\leq n}$) as long as possible.
  2. Apply a generating rule (R$\exists$C, R$\exists_{\geq n}$) and restart with step 1 as long as possible.

[3] Without loss of generality we only need to consider the case that $m = n + 1$.

In the following we always assume that rules are applied in accordance to this strategy. It ensures that the rules are applied to new individuals w.r.t. the ordering '$\prec$'.

Definition 15 (Clash Triggers) We assume the same naming conventions as used above. An ABox $\mathcal{A}$ is called *contradictory* if one of the following *clash triggers* is applicable. If none of the clash triggers is applicable to $\mathcal{A}$, then $\mathcal{A}$ is called *clash-free*.

- *Primitive clash*:
  $\mathsf{a} : \bot \in \mathcal{A}$ or $\{\mathsf{a} : \mathsf{C}, \mathsf{a} : \neg\mathsf{C}\} \subseteq \mathcal{A}$, where C is a concept name.
- *Number restriction merging clash*:
  $\{\exists_{\leq n} \mathsf{R}\} \cup \{(\mathsf{a}, \mathsf{b_k}) : \mathsf{S_k} \mid \mathsf{S_k} \in \mathsf{R}^{\downarrow}, \mathsf{k} \in 1..\mathsf{m}\} \cup \{\mathsf{b_i} \not\doteq \mathsf{b_j} \mid \mathsf{i}, \mathsf{j} \in 1..\mathsf{m}, \mathsf{i} \neq \mathsf{j}\} \subseteq \mathcal{A}$ with $\mathsf{m} > \mathsf{n}$

A clash-free ABox $\mathcal{A}$ is called *complete* if no completion rule is applicable to $\mathcal{A}$. A complete ABox $\mathcal{A}'$ derived from an ABox $\mathcal{A}$ is also called a *completion* of $\mathcal{A}$. Any ABox containing a clash is obviously unsatisfiable. The purpose of the calculus is to generate a completion for an initial ABox $\mathcal{A}$ that proves the satisfiability of $\mathcal{A}$ or its unsatisfiability if no completion can be found. In the following we have to show that a model can be constructed for any complete ABox.

4.2 Decidability of the ABox Consistency Problem

The following lemma proves that whenever a generating rule has been applied to an individual a, the concept set $\sigma(\cdot, \mathsf{a})$ of a does not change for succeeding ABoxes.

Lemma 16 (Stability) Let $\mathcal{A}$ be an ABox and $\mathsf{a} \in O_N$ be in $\mathcal{A}$. Let a generating rule be applicable to a according to the completion strategy. Let $\mathcal{A}'$ be any ABox derivable from $\mathcal{A}$ by any (possibly empty) sequence of rule applications. Then:

1. No rule is applicable in $\mathcal{A}'$ to an individual $\mathsf{b} \in O_N$ with $\mathsf{b} \prec \mathsf{a}$
2. $\sigma(\mathcal{A}, \mathsf{a}) = \sigma(\mathcal{A}', \mathsf{a})$, i.e. the concept set of a remains unchanged in $\mathcal{A}'$.
3. If $\mathsf{b} \in O_N$ is in $\mathcal{A}$ with $\mathsf{b} \prec \mathsf{a}$ then b is an individual in $\mathcal{A}'$, i.e. the individual b is not substituted by another individual.

Proof. **1.** By contradiction: Suppose $\mathcal{A} = \mathcal{A}_0 \rightarrow_* \cdots \rightarrow_* \mathcal{A}_n = \mathcal{A}'$, where $*$ is element of the completion rules and a rule is applicable to an individual b with $\mathsf{b} \prec \mathsf{a}$ in $\mathcal{A}'$. Then there has to exist a minimal i with $i \in 1..n$ such that this rule is also applicable in $\mathcal{A}_i$. If a rule is applicable to a in $\mathcal{A}$ then no rule is applicable to b in $\mathcal{A}$ due to our strategy. So no rule is applicable to any individual c such that $\mathsf{c} \prec \mathsf{a}$ in $\mathcal{A}_0, \ldots, \mathcal{A}_{i-1}$. It follows that from $\mathcal{A}_{i-1}$ to $\mathcal{A}_i$ a rule is applied to a or to a d such that $\mathsf{a} \prec \mathsf{d}$. Using an exhaustive case analysis of all rules we can show that no new assertion of the form $\mathsf{b} : \mathsf{C}$ or $(\mathsf{b}, \mathsf{e}) : \mathsf{R}$ can be added to $\mathcal{A}_{i-1}$. Therefore, no rule is applicable to b in $\mathcal{A}_i$. This is a contradiction to our assumption.

2. By contradiction: Suppose $\sigma(\mathcal{A}, \mathsf{a}) \neq \sigma(\mathcal{A}', \mathsf{a})$. Let b be the direct predecessor of a with $\mathsf{b} \prec \mathsf{a}$. A rule must have been applied to a and not to b because of point 1. Due to our strategy only generating rules are applicable to a that cannot add new elements to $\sigma(\cdot, \mathsf{a})$. This is an obvious contradiction.

3. This follows from point 1 and the completion strategy. □

Definition 17 (Blocking Individual) Let $\mathcal{A}$ be an ABox and $\mathsf{a}, \mathsf{b} \in O_N$ be individuals in $\mathcal{A}$. We call a the blocking individual of b if the following conditions hold:

1. $\mathsf{a} \succcurlyeq_{\mathcal{A}} \mathsf{b}$
2. $\mathsf{a} \prec \mathsf{b}$
3. $\neg \exists\, \mathsf{c}$ in $\mathcal{A} : \mathsf{c} \in O_N, \mathsf{c} \prec \mathsf{a}, \mathsf{c} \succcurlyeq_{\mathcal{A}} \mathsf{b}$.

The next lemma guarantees the uniqueness of a blocking individual for a blocked individual. This is a precondition for defining a particular interpretation from $\mathcal{A}$.

Lemma 18 Let $\mathcal{A}'$ be an ABox and a be a new individual in $\mathcal{A}'$. If a is blocked then

1. a has no direct successor and
2. a has exactly one blocking individual.

Proof. **1.** By contradiction: Suppose that a is blocked in $\mathcal{A}'$ and $(\mathsf{a}, \mathsf{b}) : \mathsf{R} \in \mathcal{A}'$. There must exist an ancestor ABox $\mathcal{A}$ where a generating rule has been applied to a in $\mathcal{A}$. It follows from the definition of the generating rules that for every new individual c with $\mathsf{c} \prec \mathsf{a}$ in $\mathcal{A}$ we had $\sigma(\mathcal{A}, \mathsf{c}) \not\supseteq \sigma(\mathcal{A}, \mathsf{a})$. Since $\mathcal{A}'$ has been derived from $\mathcal{A}$ we can use Lemma 16 and conclude that for every new individual c with $\mathsf{c} \prec \mathsf{a}$ in $\mathcal{A}'$ we also have $\sigma(\mathcal{A}', \mathsf{c}) \not\supseteq \sigma(\mathcal{A}', \mathsf{a})$. Thus there cannot exist a blocking individual c for a in $\mathcal{A}'$. This is a contradiction to our hypothesis.

2. This follows directly from condition 3 in Definition 17. □

Definition 19 Let A be an ABox. We define the *canonical interpretation* $\mathcal{I}_\mathcal{A} = (\Delta^{\mathcal{I}_\mathcal{A}}, \cdot^{\mathcal{I}_\mathcal{A}})$ as follows:

1. $\Delta^{\mathcal{I}_\mathcal{A}} := \{\mathsf{a} \,|\, \mathsf{a}$ is an individual in $\mathcal{A}\}$
2. $\mathsf{a}^{\mathcal{I}_\mathcal{A}} := \mathsf{a}$ iff a is mentioned in $\mathcal{A}$
3. $\mathsf{a} \in \mathsf{A}^{\mathcal{I}_\mathcal{A}}$ iff $\mathsf{a}\!:\!\mathsf{A} \in \mathcal{A}$
4. $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$ iff
   (a) $(\mathsf{a}, \mathsf{b})\!:\!\mathsf{S} \in \mathcal{A}$ for a role $\mathsf{S} \in \mathsf{R}^{\downarrow}$ *or*
   (b) $\exists\, \mathsf{c}_1, \ldots, \mathsf{c}_{\mathsf{n}-1}$ in $\mathcal{A}$: $(\mathsf{a}, \mathsf{c}_1)\!:\!\mathsf{S}_1, (\mathsf{c}_1, \mathsf{c}_2)\!:\!\mathsf{S}_2, \ldots, (\mathsf{c}_{\mathsf{n}-1}, \mathsf{b})\!:\!\mathsf{S}_\mathsf{n} \in \mathcal{A}$, $\mathsf{n} > 1$, $\mathsf{S}_\mathsf{i} \in \mathsf{R}^{\downarrow}$ for $\mathsf{i} \in 1..\mathsf{n}$ and $\mathsf{R} \in T$, *or*
   (c) $\exists\, \mathsf{c}$ in $\mathcal{A}$, $\mathsf{c} \in O_N$, c is a blocking individual for a, and $(\mathsf{c}, \mathsf{b})\!:\!\mathsf{S} \in \mathcal{A}$, for a role $\mathsf{S} \in \mathsf{R}^{\downarrow}$, *or*
   (d) $\exists\, \mathsf{c}$ in $\mathcal{A}$, $\mathsf{c} \in O_N$, c is a blocking individual for a, and $(\mathsf{c}, \mathsf{b}_1)\!:\!\mathsf{S}_1 \in \mathcal{A}$, and $\exists\, \mathsf{b}_2, \ldots, \mathsf{b}_{\mathsf{n}-1}$ in $\mathcal{A}$: $(\mathsf{b}_1, \mathsf{b}_2)\!:\!\mathsf{S}_2, \ldots, (\mathsf{b}_{\mathsf{n}-1}, \mathsf{b})\!:\!\mathsf{S}_\mathsf{n} \in \mathcal{A}$, $\mathsf{n} > 1$, $\mathsf{S}_\mathsf{i} \in \mathsf{R}^{\downarrow}$ for $\mathsf{i} \in 1..\mathsf{n}$ and $\mathsf{R} \in T$.

Theorem 20 (Soundness) Let $\mathcal{A}$ be a complete ABox, then $\mathcal{A}$ is satisfiable.

Proof. Let $\mathcal{I}_\mathcal{A} = (\Delta^{\mathcal{I}_\mathcal{A}}, \cdot^{\mathcal{I}_\mathcal{A}})$ be the canonical interpretation for the ABox $\mathcal{A}$. In the following we prove that $\mathcal{I}_\mathcal{A}$ satisfies every assertion in $\mathcal{A}$.

For any $(\mathsf{a}, \mathsf{b})\!:\!\mathsf{R} \in \mathcal{A}$ or $\mathsf{a} \dot{\neq} \mathsf{b} \in \mathcal{A}$, $\mathcal{I}_\mathcal{A}$ satisfies them by definition. Next we consider assertions of the form $\mathsf{a}\!:\!\mathsf{C}$. We show by induction on the structure of C that $\mathsf{a} \in \mathsf{C}^{\mathcal{I}_\mathcal{A}}$.

If C is a concept name, then $\mathsf{a} \in \mathsf{C}^{\mathcal{I}_\mathcal{A}}$ by definition of $\mathcal{I}_\mathcal{A}$. If $\mathsf{C} = \top$, then obviously $\mathsf{a} \in \top^{\mathcal{I}_\mathcal{A}}$. The case $\mathsf{C} = \bot$ cannot occur since $\mathcal{A}$ is clash-free.

If $\mathsf{C} = \neg\mathsf{D}$, then D is a concept name since all concepts are in negation normal form (see Definition 9). $\mathcal{A}$ is clash-free and cannot contain $\mathsf{a}\!:\!\mathsf{D}$. Thus, $\mathsf{a} \notin \mathsf{D}^{\mathcal{I}_\mathcal{A}}$, i.e. $\mathsf{a} \in \Delta^{\mathcal{I}_\mathcal{A}} \setminus \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. Hence $\mathsf{a} \in (\neg\mathsf{D})^{\mathcal{I}_\mathcal{A}}$.

If $\mathsf{C} = \mathsf{C}_1 \sqcap \mathsf{C}_2$ then (since $\mathcal{A}$ is complete) $\mathsf{a}\!:\!\mathsf{C}_1 \in \mathcal{A}$ and $\mathsf{a}\!:\!\mathsf{C}_2 \in \mathcal{A}$. By induction hypothesis, $\mathsf{a} \in {\mathsf{C}_1}^{\mathcal{I}_\mathcal{A}}$ and $\mathsf{a} \in {\mathsf{C}_2}^{\mathcal{I}_\mathcal{A}}$. Hence $\mathsf{a} \in (\mathsf{C}_1 \sqcap \mathsf{C}_2)^{\mathcal{I}_\mathcal{A}}$.

If $\mathsf{C} = \mathsf{C}_1 \sqcup \mathsf{C}_2$ then (since $\mathcal{A}$ is complete) either $\mathsf{a}\!:\!\mathsf{C}_1 \in \mathcal{A}$ or $\mathsf{a}\!:\!\mathsf{C}_2 \in \mathcal{A}$. By induction hypothesis, $\mathsf{a} \in {\mathsf{C}_1}^{\mathcal{I}_\mathcal{A}}$ or $\mathsf{a} \in {\mathsf{C}_2}^{\mathcal{I}_\mathcal{A}}$. Hence $\mathsf{a} \in (\mathsf{C}_1 \sqcup \mathsf{C}_2)^{\mathcal{I}_\mathcal{A}}$.

If $\mathsf{C} = \forall\, \mathsf{R} \,.\, \mathsf{D}$, then we have to show that for all b with $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$ it holds that $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. If $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$, then according to Definition 19 the following cases can occur: (4a) b is a direct S-successor of a for a role $\mathsf{S} \in \mathsf{R}^{\downarrow}$ with $\mathsf{S}^{\mathcal{I}_\mathcal{A}} \subseteq \mathsf{R}^{\mathcal{I}_\mathcal{A}}$; then we have $\mathsf{b}\!:\!\mathsf{D} \in \mathcal{A}$ since $\mathcal{A}$ is complete and by induction hypothesis $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. (4b) b is a R-successor of a via a subrole chain of S_i's with $\mathsf{S_i}^{\mathcal{I}_\mathcal{A}} \subseteq \mathsf{R}^{\mathcal{I}_\mathcal{A}}, \mathsf{R} \in T$; then we have $\mathsf{c_{n-1}}\!:\!\forall\, \mathsf{R} \,.\, \mathsf{D} \in \mathcal{A}$ and $\mathsf{b}\!:\!\mathsf{D} \in \mathcal{A}$ since $\mathcal{A}$ is complete and by induction hypothesis we have $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. (4c) There has to exist a blocking individual c such that $\mathsf{c}\!:\!\forall\, \mathsf{R} \,.\, \mathsf{D} \in \mathcal{A}$ and $(\mathsf{c}, \mathsf{b})\!:\!\mathsf{S} \in \mathcal{A}$ for a role $\mathsf{S} \in \mathsf{R}^{\downarrow}$ and because $\mathcal{A}$ is complete we have $\mathsf{b}\!:\!\mathsf{D} \in \mathcal{A}$ and again by induction hypothesis it holds $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. (4d) This case combines the cases (4b-c) because the individual b is reachable from the blocking individual c via a chain of subroles of the transitive role R. It can be proven analogously.

If $\mathsf{C} = \exists\, \mathsf{R} \,.\, \mathsf{D}$, then we have to show that there exists an individual $\mathsf{b} \in \Delta^{\mathcal{I}_\mathcal{A}}$ with $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$ and $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. Since ABox $\mathcal{A}$ is complete, we have either $(\mathsf{a}, \mathsf{b})\!:\!\mathsf{R} \in \mathcal{A}$ and $\mathsf{b}\!:\!\mathsf{D} \in \mathcal{A}$ or a is blocked by an individual c and $(\mathsf{c}, \mathsf{b})\!:\!\mathsf{R} \in \mathcal{A}$. In the first case we have $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$ and $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$ by induction hypothesis and the definition of $\mathcal{I}_\mathcal{A}$. In the second case there exists the blocking individual c with $\mathsf{c}\!:\!\exists\, \mathsf{R} \,.\, \mathsf{D} \in \mathcal{A}$. By definition c cannot be blocked and by hypothesis $\mathcal{A}$ is complete. So we have an individual b with $(\mathsf{c}, \mathsf{b})\!:\!\mathsf{R} \in \mathcal{A}$ and $\mathsf{b}\!:\!\mathsf{D} \in \mathcal{A}$. By induction hypothesis we have $\mathsf{b} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$ and by the definition of $\mathcal{I}_\mathcal{A}$ (case 4c) we have $(\mathsf{a}, \mathsf{b}) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$.

If $\mathsf{C} = \exists_{\geq n}\, \mathsf{R}$, we prove the hypothesis by contradiction. We assume that $\mathsf{a} \notin (\exists_{\geq n}\, \mathsf{R})^{\mathcal{I}_\mathcal{A}}$. Then there exist at most m $(0 \leq m < n)$ distinct R-successors of a. Two cases can occur: (1) the individual a is not blocked in $\mathcal{I}_\mathcal{A}$. Then we have less than n R-successors of a in $\mathcal{A}$ and the $\mathrm{R}\exists_{\geq n}$-rule is applicable to a. This contradicts the assumption that $\mathcal{A}$ is complete. (2) a is blocked by an individual c but the same argument as in case (1) holds and leads to the same contradiction.

For $\mathsf{C} = \exists_{\leq n}\, \mathsf{R}$ we show the goal by contradiction. Suppose that $\mathsf{a} \notin (\exists_{\leq n}\, \mathsf{R})^{\mathcal{I}_\mathcal{A}}$. Then there exist at least $n+1$ distinct individuals $\mathsf{b}_1, \ldots, \mathsf{b}_{n+1}$ such that $(\mathsf{a}, \mathsf{b}_i) \in \mathsf{R}^{\mathcal{I}_\mathcal{A}}$, $i \in 1..n+1$. According to Definition 19 the following two cases can occur. (1) We have $n+1$ $(\mathsf{a}, \mathsf{b}_i)\!:\!\mathsf{S}_i \in \mathcal{A}$ with $\mathsf{S}_i \in \mathsf{R}^{\downarrow}$ and $\mathsf{S}_i \notin T$, $i \in 1..n+1$. The $\mathrm{R}\exists_{\leq n}$ rule cannot be applicable since $\mathcal{A}$ is complete and the b_i are distinct, i.e. $\mathsf{b}_i \not\doteq \mathsf{b}_j \in \mathcal{A}$, $i, j \in 1..n+1$, $i \neq j$. This contradicts the assumption that $\mathcal{A}$ is clash-free. (2) There exists a blocking individual c with $(\mathsf{c}, \mathsf{b}_i)\!:\!\mathsf{S}_i \in \mathcal{A}$, $\mathsf{S}_i \in \mathsf{R}^{\downarrow}$, and $\mathsf{S}_i \notin T$, $i \in 1..n+1$. This leads to an analogous contradiction.

If $\forall\, x \,.\, x\!:\!\mathsf{D} \in \mathcal{A}$, then –due to the completeness of $\mathcal{A}$– for each individual a in $\mathcal{A}$ we have $\mathsf{a}\!:\!\mathsf{D} \in \mathcal{A}$ and, by the previous cases, $\mathsf{a} \in \mathsf{D}^{\mathcal{I}_\mathcal{A}}$. Thus, $\mathcal{I}_\mathcal{A}$ satisfies $\forall\, x \,.\, x\!:\!\mathsf{D}$. Finally, since $\mathcal{I}_\mathcal{A}$ satisfies all assertions in $\mathcal{A}$, $\mathcal{I}_\mathcal{A}$ satisfies $\mathcal{A}$. □

Theorem 21 (Completeness) Let $\mathcal{A}$ be a satisfiable ABox, then there exists at least one completion of $\mathcal{A}$ computed by applying the completion rules.

Proof. Obviously, an Abox containing a clash is unsatisfiable. If every completion of $\mathcal{A}$ is unsatisfiable, then it follows from Proposition 12 that ABox

$\mathcal{A}$ is unsatisfiable. □

Definition 22 For any augmentation of an initial ABox $\mathcal{A}$, we define the *concept size* $n_\mathcal{A}$ as the number of concepts or subconcepts occurring in $\mathcal{A}$.[4] Note that $n_\mathcal{A}$ is bound by the length of the string expressing $\mathcal{A}$. The *size* of an ABox $\mathcal{A}$ is defined as $n_\mathcal{A} \times \| T \| + \| O_O \|$.

Lemma 23 Let $\mathcal{A}$ be an ABox and let $\mathcal{A}'$ be a completion of $\mathcal{A}$. In any set X consisting of individuals occurring in $\mathcal{A}'$ with a cardinality greater than $2^{n_\mathcal{A}}$ there exist at least two individuals $\mathsf{a}, \mathsf{b} \in X$ whose concept sets are equal ($\mathsf{a} \equiv_{\mathcal{A}'} \mathsf{b}$).

Proof. Each assertion $\mathsf{a} : \mathsf{C_i} \in \mathcal{A}'$ may contain at most $n_\mathcal{A}$ different concepts $\mathsf{C_i}$. So there cannot exist more than $2^{n_\mathcal{A}}$ different concept sets for the individuals in $\mathcal{A}'$. □

Lemma 24 Let $\mathcal{A}$ be an ABox and let $\mathcal{A}'$ be a completion of $\mathcal{A}$. Then there occur at most $2^{n_\mathcal{A}}$ non-blocked new individuals in $\mathcal{A}'$.

Proof. Suppose we have $2^{n_\mathcal{A}} + 1$ non-blocked new individuals in $\mathcal{A}'$. From Lemma 23 we know that there exist at least two individuals a, b in $\mathcal{A}'$ such that $\mathsf{a} \equiv_{\mathcal{A}'} \mathsf{b}$. By Definition 13 we have either $\mathsf{a} \prec \mathsf{b}$ or $\mathsf{b} \prec \mathsf{a}$. Assume without loss of generality that $\mathsf{a} \prec \mathsf{b}$ holds and $\mathsf{a} \equiv_{\mathcal{A}'} \mathsf{b}$ implies $\sigma(\mathcal{A}', \mathsf{a}) \supseteq \sigma(\mathcal{A}', \mathsf{b})$. Then we have either $\mathsf{a} \succcurlyeq_{\mathcal{A}'} \mathsf{b}$ or there exists an individual c with $\mathsf{c} \succcurlyeq_{\mathcal{A}'} \mathsf{b}$ and $\mathsf{c} \prec \mathsf{a}$. Both cases contradict the hypothesis. □

Theorem 25 (Termination) Let $\mathcal{A}_\mathcal{T}$ be the augmented ABox w.r.t a TBox $\mathcal{T}$ and let n be the size of $\mathcal{A}_\mathcal{T}$. Every completion of $\mathcal{A}_\mathcal{T}$ is finite and its size is $O(2^{4n})$.

Proof. Let $\mathcal{A}'$ be a completion of $\mathcal{A}_\mathcal{T}$. From Lemma 24 we know that $\mathcal{A}'$ has at most 2^n non-blocked new individuals. Therefore, a total of at most $m \times 2^n$ new individuals may exists in $\mathcal{A}'$, where m is the maximum number of direct successors for any individual in $\mathcal{A}'$.

Note that m is bound by the number of $\exists\, \mathsf{R} \,.\, \mathsf{C}$ concepts ($\leq n$) plus the total sum of numbers occurring in $\exists_{\geq n}\, \mathsf{R}$. Since numbers are expressed in binary, their sum is bound by 2^n. Hence, we have $m \leq 2^n + n$. Since the number of individuals in the initial ABox is also bound by n, the total number of individuals in $\mathcal{A}'$ is at most $m \times (2^n + n) \leq (2^n + n) \times (2^n + n)$, i.e. $O(2^{2n})$.

The number of different assertions of the form $\mathsf{a} : \mathsf{C}$ or $\forall\, x \,.\, x : \mathsf{C}$ in which each individual in $\mathcal{A}'$ can be involved, is bound by n and each assertion has a size

[4]We have to increase $n_\mathcal{A}$ by 1 if $\top$ does not occur in $\mathcal{A}$.

linear in n. Hence, the total size of these assertions is bound $n \times n \times 2^{2n}$, i.e. $O(2^{3n})$.

The number of different assertions of the form $(\mathsf{a}, \mathsf{b}) : \mathsf{R}$ or $\mathsf{a} \neq \mathsf{b}$ is bound by $(2^{2n})^2$, i.e. $O(2^{4n})$. In conclusion, we have a size of $O(2^{4n})$ for $\mathcal{A}'$. □

Theorem 26 (Decidability) Let $\mathcal{A}_{\mathcal{T}}$ be an ABox w.r.t. a TBox $\mathcal{T}$. Checking whether $\mathcal{A}_{\mathcal{T}}$ is satisfiable is a decidable problem.

Proof. This follows immediately from the Theorems 20, 21, and 25. □

5 Conclusion

We presented the first treatment for a tableaux calculus deciding the ABox consistency problem for the description logic $\mathcal{ALCNH}_{R^+}$. A highly optimized variant of this calculus is already implemented in the ABox description logic system RACE[5] [Haarslev et al., 1999] demonstrating the practical usefulness of $\mathcal{ALCNH}_{R^+}$. Although TBox reasoners for logics such as $\mathcal{ALCQHI}_{R^+}$ are available, the development of $\mathcal{ALCNH}_{R^+}$ and its optimized implementation in RACE is a novel approach. Practical reasoning is only possible with the design and implementation of appropriate optimization techniques. This is supported by recent empirical findings [Haarslev and Möller, 2000] suggesting that RACE dramatically outperforms other known DL reasoners for logics at least as expressive as $\mathcal{ALCNH}_{R^+}$. To the best of our knowledge there currently exists no other ABox DL system with a performance comparable to RACE.

Acknowledgements

We are grateful to Carsten Lutz, Stephan Tobies, and Michael Wessel for detailed and thoughtful comments on an earlier version of this paper. Recently, in a parallel development, a calculus for deciding the ABox consistency problem for the description logic $\mathcal{ALCQHI}_{R^+}$ has been proposed in a revised version of [Horrocks et al., 1999a].

References

Buchheit, M., Donini, F., and Schaerf, A. (1993). Decidable reasoning in terminological knowledge representation systems. *Journal of Artificial Intelligence Research*, 1:109–138.

De Giacomo, G. and Lenzerini, M. (1996). TBox and ABox reasoning in expressive description logics. In Aiello, L., Doyle, J., and Shapiro, S.,

[5]Available from http://kogs-www.informatik.uni-hamburg.de/~moeller/race.html

editors, *Fifth International Conference on Principles of Knowledge Representation, Cambridge, Mass., Nov. 5-8, 1996*.

Haarslev, V. and Möller, R. (2000). Optimization techniques for reasoning with expressive ABox and concept expressions. Technical Report FBI-HH-M-290/00, University of Hamburg, Computer Science Department. Available at URL http://kogs-www.informatik.uni-hamburg.de/~haarslev/publications/report-FBI-290-2000.ps.gz.

Haarslev, V., Möller, R., and Turhan, A.-Y. (1999). RACE user's guide and reference manual version 1.1. Technical Report FBI-HH-M-289/99, University of Hamburg, Computer Science Department. Available at URL http://kogs-www.informatik.uni-hamburg.de/~haarslev/publications/report-FBI-289-99.ps.gz.

Hollunder, B. (1994). *Algorithmic Foundations of Terminological Knowledge Representation Systems*. PhD thesis, University of Saarbrücken, Department of Computer Science.

Horrocks, I. (1998). Using an expressive description logic: FaCT or fiction? In Cohn, T., Schubert, L., and Shapiro, S., editors, *Proceedings of Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98), Trento, Italy, June 2-5, 1998*, pages 636–647.

Horrocks, I. and Patel-Schneider, P. (1999). Optimising description logic subsumption. *Journal of Logic and Computation*, 9(3):267–293.

Horrocks, I. and Sattler, U. (1999). A description logic with transitive and inverse roles and role hierarchies. *Journal of Logic and Computation*, 9(3):385–410.

Horrocks, I., Sattler, U., and Tobies, S. (1999a). A description logic with transitive and converse roles, role hierarchies and qualifying number restrictions. Technical Report LTCS-99-08, RWTH Aachen.

Horrocks, I., Sattler, U., and Tobies, S. (1999b). Practical reasoning for expressive description logics. In Ganzinger, H., McAllester, D., and Voronkov, A., editors, *Proceedings of the 6th International Conference on Logic for Programming and Automated Reasoning (LPAR'99)*, number 1705 in Lecture Notes in Artificial Intelligence, pages 161–180. Springer-Verlag.

Lewis, H. and Papadimitriou, C. (1981). *Elements of the Theory of Computation*. Prentice-Hall, Englewood Cliffs, New Jersey.

Sattler, U. (1996). A concept language extended with different kinds of transitive roles. In Görz, G. and Hölldobler, S., editors, *20. Deutsche Jahrestagung für Künstliche Intelligenz*, number 1137 in Lecture Notes in Artificial Intelligence, pages 333–345. Springer Verlag, Berlin.

Schmidt-Schauss, M. and Smolka, G. (1991). Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26.

Tessaris, S. and Gough, G. (1999). ABox reasoning with transitive roles and axioms. In Lambrix et al., P., editor, *Proceedings of the International Workshop on Description Logics (DL'99), July 30 - August 1, 1999, Linköping, Sweden*, pages 101–104.