In *Associazione italiana per l'intelligenza artificiale*, 5(2), 1992.

On the Tractability of Terminological Logics with Numerical Restrictions

Fabrizio Sebastiani & Umberto Straccia
Istituto di Elaborazione dell'Informazione
Consiglio Nazionale delle Ricerche
Via S. Maria, 46 - 56126 Pisa (Italy)
e-mail: fabrizio@icnucevm.cnuce.cnr.it

Abstract

A number of results relative to the complexity of terminological logics have recently appeared in the literature. Unfortunately, most of these results are "negative", as they show that, in the logics they refer to, deciding *subsumption* is intractable. In this paper we show that computing subsumption is $O(n^2)$ in $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$, a logic obtained by adding the two operators **atleast** and **atmost**, which allow the specification of number restrictions, to Brachman and Levesque's $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{-}$ logic.

1 Introduction

There is a well-known tradeoff between the expressive power of a logic and its computational tractability. One of the fundamental problems of the research on *terminological logics* (TLs) is that of individuating TLs which are "optimal" from the standpoint of this tradeoff, i.e. those TLs which are most expressive among the ones for which a polynomial decision algorithm can be given. A number of results relative to the computational complexity of the decision problem in TLs have recently appeared in the literature. The best known result in this sense is the one due to Brachman and Levesque [1], who have shown that deciding *subsumption* (i.e. the metalinguistic relation analogous to the one of validity in standard logics) is co-NP-hard for $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle = \{$**and**, **all**, **some**, **restr**$\}$ and $O(n^2)$ for $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{-} = \mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle \setminus \{$**restr**$\}$. Unfortunately, nearly all results appeared after this one are "negative" in nature, as they concern either the undecidability or the computational intractability of a number of TLs.

In this work we show that deciding subsumption is $O(n^2)$ in the logic $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$, an extension of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{-}$ obtained by the addition of the two operators **atleast** and **atmost**. This result also confirms a hypothesis by von Luck and Owsnicki-Klewe [5], who had conjectured that $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$ had polynomial complexity.[1]

2 Syntax and semantics of the $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$ logic

The $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$ logic, like many other TLs, allows the specification of two fundamental types of terms: *concepts* and *roles*. Concepts are terms denoting sets of individuals, and are, so to speak, the first-class citizens of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$; in fact, it is only on concepts that the subsumption relation is defined. Roles are terms denoting binary relations between individuals; their function is to allow the specification of structural constituents of concepts.

The syntax of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$ is the following:

< *concept* > ::= < *atom* >
| (**and** < *concept* >$^+$)
| (**all** < *role* > < *concept* >)
| (**atleast** < *positive integer* > < *role* >)
| (**atmost** < *positive integer* > < *role* >)
< *role* > ::= < *atom* >

[1] This result has independently been shown also by Donini *et al.* [2]; their result is actually stronger than ours, in that they give a tractability result for a logic that strictly contains $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$. The methods for proving the result are different in the two cases; Donini and colleagues use the constraint-propagation method of [3], while we have used a method structurally similar to the one used in [1].

The formal semantics of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$ is detailed in the following definition.

Definition 1
Let $\mathcal{D}$ be a nonempty set of individuals and $\mathcal{E}$ a function from concepts into subsets of $\mathcal{D}$ and from roles into subsets of $\mathcal{D} \times \mathcal{D}$. $\mathcal{E}$ is an **extension function** *over $\mathcal{D}$ if and only if:*

1. $\mathcal{E}[\![(\textbf{and } F_1 \ldots\ F_n)]\!] \ = \ \bigcap_{i=1}^{n} \mathcal{E}[\![F_i]\!]$
2. $\mathcal{E}[\![(\textbf{all } R\ F)]\!] \ = \ \{x \in \mathcal{D} | \forall y < x, y > \in \mathcal{E}[\![R]\!] \Rightarrow y \in \mathcal{E}[\![F]\!]\}$
3. $\mathcal{E}[\![(\textbf{atleast } n\ R)]\!] \ = \ \{x \in \mathcal{D} |\ \parallel \{y \in \mathcal{D} |\ < x, y > \in \mathcal{E}[\![R]\!]\} \parallel \geq n\}$
4. $\mathcal{E}[\![(\textbf{atmost } n\ R)]\!] \ = \ \{x \in \mathcal{D} |\ \parallel \{y \in \mathcal{D} |\ < x, y > \in \mathcal{E}[\![R]\!]\} \parallel \leq n\}$

The notion of subsumption between concepts is then specified by means of the following definition.

Definition 2
Let F_1 and F_2 be two concepts of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$. F_2 is **subsumed** *by F_1 if and only if for every nonempty set of individuals $\mathcal{D}$ and every extension function $\mathcal{E}$ over $\mathcal{D}$ it is true that $\mathcal{E}[\![F_2]\!] \subseteq \mathcal{E}[\![F_1]\!]$.*

3 The computational tractability of $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$

In this section we will argue that there exists an algorithm which decides in time $O(n^2)$ if a concept F_1 subsumes a concept F_2 in $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$. The algorithm is essentially divided in two parts: the first part reduces a concept into a "normal form", while the second part checks whether there exists a relation of subsumption between two already normalized concepts. The algorithm is structurally similar to the subsumption algorithm given for $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{-}$ in [1], but is a non-trivial extension of it, as non-trivial is the proof of its completeness wrt $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$.

In the full paper [4] we define the notion of *normal form* of a concept and describe a normalization algorithm NORM; we then prove the following lemma that describes its properties.

Lemma 1 *Normalization*
Let F be a concept of $\mathcal{B}\nabla\dashv\rfloor\langle\updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$. The NORM algorithm reduces, in time $O(|F|^2)$, the concept F in a normal form F' such that $\mathcal{E}[\![F]\!] = \mathcal{E}[\![F']\!]$ for any extension function $\mathcal{E}$, and such that $|F'|$ is $O(|F|)$. ∎

Normalization structures a concept in a tree-like form, with **all** operators as internal nodes and all other constituents as leaves. Normalization preserves then both the extension and (modulo a multiplicative constant) the length of the concept being normalized.

In the full paper we describe SUBS, an algorithm which takes as input two concepts in normal form and decides whether a relation of subsumption exists between them; we then go on to prove the following lemma.

Lemma 2 *Complexity*
The SUBS(F_1, F_2) *algorithm runs in time $O(n^2)$, where $n = max\{|F_1|, |F_2|\}$.*
∎

Given the $O(n^2)$ result for the normalization of concepts F_1 and F_2, we may conclude that the complexity of computing subsumption between two not yet normalized concepts is also $O(n^2)$.

The following lemma is easily proven by induction on the maximal depth of the two concepts involved.

Lemma 3 *Correctness*
If SUBS(F_1, F_2)=*true, then F_2 is subsumed by F_1.* ∎

Let us now switch to the completeness proof for SUBS. In the full paper we describe the BUILD-EXTENSION+ algorithm which, given a concept F in normal form, a set of individuals Φ and an extension function $\mathcal{E}$ defined over a domain $\mathcal{D}$ such that $\Phi \cap \mathcal{D} = \emptyset$, "expands" $\mathcal{E}$ by building an extension function $\mathcal{E}'$ defined over a domain $\mathcal{D}'$ such that $\Phi \subseteq \mathcal{D}'$ and $\Phi \subseteq \mathcal{E}'[\![F]\!]$, and such that for every nonempty set $\Psi \subseteq \mathcal{D}$ it turns out that $\Psi \subseteq \mathcal{E}[\![F]\!]$ if and only if $\Psi \subseteq \mathcal{E}'[\![F]\!]$. This algorithm will be useful in determining that, if SUBS returns *false*, there exists an extension function $\mathcal{E}$ and a set of individuals Φ such that $\Phi \subseteq \mathcal{E}[\![F_2]\!]$ and $\Phi \not\subseteq \mathcal{E}[\![F_1]\!]$, and from this that F_2 is not subsumed by F_1.

The BUILD-EXTENSION+ plays a critical role in the proof of the following lemma.

Lemma 4 *Existence*
*Let F=(**and** $a_1 \ldots a_n$) be a concept in normal form, Φ and Ψ two sets of individuals, $\Psi \subseteq \mathcal{D}$ a nonempty set of individuals, $\mathcal{E}$ an extension function over $\mathcal{D}$ such that $\Phi \cap \mathcal{D} = \emptyset$ and $\mathcal{E}'$ the extension function defined over the domain $\mathcal{D}'$ which is returned by the procedure* BUILD-EXTENSION+(F,Φ,$\mathcal{E}$,$\mathcal{D}$). *The following conditions hold:*

- $\Phi \subseteq \mathcal{D}'$
- $\Phi \subseteq \mathcal{E}'[\![F]\!]$;
- $\Psi \subseteq \mathcal{E}[\![F]\!]$ *if and only if* $\Psi \subseteq \mathcal{E}'[\![F]\!]$. ∎

We now switch to the proof of completeness.

Lemma 5 *Completeness*
Let F_1 and F_2 be two concepts in normal form. If SUBS(F_1, F_2)=false, *then F_2 is not subsumed by F_1.* ∎

The proof of this lemma is rather complex, and is described in detail only in the full paper. The proof proceeds by induction on the maximal depth k of F_1 and F_2; by using the existence lemma it may be shown that, when SUBS(F_1, F_2) returns *false*, there exists an extension function $\mathcal{E}$ and a nonempty set of individuals Φ such that $\Phi \subseteq \mathcal{E}[\![F_2]\!]$ e $\Phi \not\subseteq \mathcal{E}[\![F_1]\!]$.

Given the normalization, complexity, correctness and completeness lemmas, the following theorem follows.

Theorem 1
Given two concepts F_1 and F_2 in $\mathcal{B}\nabla\dashv\rfloor\langle\Updownarrow\dashv\backslash\forall\triangle^{\mathcal{N}-}$, it may be determined in $O(n^2)$ if F_2 is subsumed by F_1, where $n = max(\{|F_1|, |F_2|\})$. ∎

References

[1] R.J. Brachman and H.J. Levesque. The Tractability of Subsumption in Frame-Based Description Languages. In *Proceedings of AAAI-84*, pages 34-37, Austin, TX, 1984.

[2] F. Donini, B. Hollunder, M. Lenzerini, D. Nardi and W. Nutt. Tractable terminological languages. This volume.

[3] M. Schmidt-Schauß and G. Smolka. *Attributive Concept Descriptions with Unions and Complements*. IWBS Report 68, IBM Germany Scientific Center, Stuttgart, FRG, 1989.

[4] F. Sebastiani and U. Straccia. *A Computationally Tractable Terminological Logic*. Unpublished manuscript.

[5] K. von Luck and B. Owsnicki-Klewe. *New AI Formalisms for Knowledge Representation: a Case Study*. KIT Report 41, Fachbereich Informatik, Technische Universität Berlin, Berlin, 1987.