Uniform Interpolation for $\mathcal{ALC}$ Revisited

Zhe Wang$^1$, Kewen Wang$^1$, Rodney Topor$^1$, Jeff Z. Pan$^2$, and Grigoris Antoniou$^3$

$^1$ Griffith University, Australia
$^2$ University of Aberdeen, UK
$^3$ University of Crete, Greece

Abstract. The notion of uniform interpolation for description logic $\mathcal{ALC}$ has been introduced in [9]. In this paper, we reformulate the uniform interpolation for $\mathcal{ALC}$ from the angle of forgetting and show that it satisfies all desired properties of forgetting. Then we introduce an algorithm for computing the result of forgetting in concept descriptions. We present a detailed proof for the correctness of our algorithm using the Tableau for $\mathcal{ALC}$. Our results have been used to compute forgetting for $\mathcal{ALC}$ knowledge bases.

1 Introduction

The Web Ontology Language (OWL) [15] provides a construct owl:imports for importing and merging Web ontologies by referencing axioms contained in another ontology that may be located somewhere else on the Web. This construct is very limited in that it can only merge some linked ontologies together but is unable to resolve conflicts among merged ontologies or to filter redundant parts from those ontologies. However, an ontology is often represented as a logical theory, and the removal of one term may influence other terms in the ontology. Thus, more advanced methods for dealing with large ontologies and reusing existing ontologies are desired.

It is well-known that OWL is built on description logics (DLs) [11]. Recent efforts show that the notions of uniform interpolation and forgetting are promising tools for extracting modular ontologies from a large ontology. In a recent experiment reported in [5], uniform interpolation and forgetting have been used for extracting modular ontologies from two large medical ontologies SNOMED CT [16] and NCI [17]. SNOMED CT contains around 375,000 concept definitions while NCI Thesaurus has 60,000 axioms. The experiment result is promising. For instance, if 2,000 concepts definitions are forgotten from SNOMED CT, the success rate is 93% and if 5,000 concepts definitions are forgotten from NCI, the success rate is 97%.

Originally, interpolation is proposed and investigated in pure mathematical logic, specifically, in proof theory. Given a theory $T$, ordinary interpolation for $T$ says that if $T \vdash \phi \rightarrow \psi$ for two formulas $\phi$ and $\psi$, then there is a formula $I(\phi, \psi)$ in the language containing only the shared symbols, say $S$, such that $T \vdash \phi \rightarrow I(\phi, \psi)$ and $T \vdash I(\phi, \psi) \rightarrow \psi$. Uniform interpolation is a strengthening of ordinary interpolation in that the interpolant can be obtained from either $\phi$ and $S$ or from $\psi$ and $S$. Uniform

* This work was partially supported by the Australia Research Council (ARC) Discovery Project 0666107.

A. Nicholson and X. Li (Eds.): AI 2009, LNAI 5866, pp. 528–537, 2009.
© Springer-Verlag Berlin Heidelberg 2009
interpolation for various propositional modal logics have been investigated by Visser [10] and Ghilardi [3]. A definition of uniform interpolation for the description logic $\mathcal{ALC}$ is given in [9] and it is used in investigating the definability of TBoxes for $\mathcal{ALC}$.

On the other hand, (semantic) forgetting is studied by researchers in AI [8,7,2]. Informally, given a knowledge base $K$ in classical logic or nonmonotonic logic, we may wish to forget about (or discard) some redundant predicates but still preserve certain forms of reasoning. Forgetting has been investigated for DL-Lite and extended $\mathcal{EL}$ in [11,6,5] but not for expressive DLs. Forgetting for modal logic is studied in [13].

Forgetting and uniform interpolation have different intuitions behind them and are introduced by different communities. Uniform interpolation is originally investigated as a syntactic concept and forgetting is a semantic one. However, if the axiom system is sound and complete, they can be characterized by each other.

In this paper, we first reformulate the notion of uniform interpolation for $\mathcal{ALC}$ studied in [9] from the angle of forgetting and show that all desired properties of forgetting are satisfied. We introduce an algorithm for computing the result of forgetting in concept descriptions and, a novel and detailed proof for the correctness of the algorithm is developed using the Tableau for $\mathcal{ALC}$. We note that a similar algorithm for uniform interpolation is provided in [9] in which it is mentioned that the correctness of their algorithm can be shown using a technique called bisimulation that is widely used in modal logic. In a separate paper [12], we use the results obtained in this paper to compute forgetting for $\mathcal{ALC}$ knowledge bases.

Due to space limitation, proofs are omitted in this paper but can be found at [http://www.cit.gu.edu.au/~kewen/Papers/alc_forget_long.pdf](http://www.cit.gu.edu.au/~kewen/Papers/alc_forget_long.pdf).

2 Preliminaries

We briefly recall some basics of $\mathcal{ALC}$. Further details of $\mathcal{ALC}$ and other DLs can be found in [1].

First, we introduce the syntax of concept descriptions for $\mathcal{ALC}$.

Elementary concept descriptions consist of both concept names and role names. So a concept name is also called atomic concept while a role name is also called atomic role. Complex concept descriptions are built inductively as follows: $A$ (atomic concept); $\top$ (universal concept); $\bot$ (empty concept); $\neg C$ (negation); $C \sqcap D$ (conjunction); $C \sqcup D$ (disjunction); $\forall R.C$ (universal quantification) and $\exists R.C$ (existential quantification). Here, $A$ is an (atomic) concept, $C$ and $D$ are concept descriptions, and $R$ is a role.

An interpretation $\mathcal{I}$ of $\mathcal{ALC}$ is a pair $(\Delta^\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta^\mathcal{I}$ is a non-empty set called the domain and $\cdot^\mathcal{I}$ is an interpretation function which associates each (atomic) concept $A$ with a subset $A^\mathcal{I}$ of $\Delta^\mathcal{I}$ and each atomic role $R$ with a binary relation $R^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. The function $\cdot^\mathcal{I}$ can be naturally extended to complex descriptions:

\[
\begin{align*}
\top^\mathcal{I} &= \Delta^\mathcal{I} \\
(\neg C)^\mathcal{I} &= \Delta^\mathcal{I} - C^\mathcal{I} \\
(C \sqcap D)^\mathcal{I} &= C^\mathcal{I} \cap D^\mathcal{I} \\
(C \sqcup D)^\mathcal{I} &= C^\mathcal{I} \cup D^\mathcal{I} \\
\bot^\mathcal{I} &= \emptyset \\
(C \sqcap D)^\mathcal{I} &= C^\mathcal{I} \cap D^\mathcal{I}
\end{align*}
\]

1 An email communication with one of the authors of [9] shows that they have not got a complete proof yet.