Satisfiability and reasoning mechanism of terminological cycles in description logic $\mathcal{L}$

WANG Ju†, JIANG YunCheng & SHEN YuMing

College of Computer Science and Information Engineering, Guangxi Normal University, Guilin 541004, China

The current research works and the existing problems of terminological cycles in description logics are analyzed in this paper. Referring to the works of Baader F and Nebel B, we aim in a new direction. Firstly, description logic $\mathcal{L}$ is defined, and the description graphs GT and GJ are redefined. A syntax condition for the satisfiability of membership relation is given. By using this syntax condition, we prove the following: The subsumption reasoning in $\mathcal{L}$ with respect to gfp-model, lfp-model and descriptive model is polynomial.

description logic, terminological cycles, description graph, model

1 Introduction

Terminological cycles (or cyclic definitions) have been a difficult spot in the study of description logics for quite a few years. The basic problems, such as their semantics and reasoning mechanisms, have not been reasonably well settled\cite{1,2}. Therefore, in the description logics reasoning systems which have been implemented, such as Pellet\cite{3}, FaCT++\cite{4}, Racer\cite{5}, and so on, there is always a mandatory regulation: description logic knowledge bases T-Box does not have terminological cycles\cite{6}. However, terminological cycles may greatly extend the expressive capability of the language. In some applications (such as medical field), terminological cycles are indispensable\cite{7}. Also, by using terminological cycles, one can establish a description logic knowledge base more conveniently because it would give the people an intuitive understanding of the axioms, sentences of the knowledge base. If we do not accept terminological cycles in the knowledge base, when a cyclic concept occurs in the application, it would make the system specification overly complex and hard to understand by a user\cite{8,9}. For these reasons, we think it is significant to study the semantics and the algorithmic nature of terminological cycles.
Let $T$ be a T-Box with terminological cycles. Let $J$ be a base interpretation (or primitive interpretation) for $T: \left\langle \Delta; B_1^j, \ldots, B_n^j; R_1^j, \ldots, R_m^j \right\rangle$, where $B_i^j \subseteq \Delta$, $R_i^j \subseteq \Delta \times \Delta$. $B_i$ and $R_i$ are called base symbols and roles, respectively. Let $N_1, N_2, \ldots, N_k$ be defined concepts (or name symbols, defined symbols in some literatures), if for each $N_i$, we assign a subset $N_i^j \subseteq \Delta$ of $\Delta$ to it; these assignments together with base interpretation $J$ (denoted by $I$) are called an extension of $J$. If for any concept axiom

$$A = T(A),$$

we always have

$$A' = (T(A))'. $$

Then, $I$ is called a model of $T$. If all $N_i^j$ are empty sets, (generally, if all $N_i^j - B_i^j \cup B_i^j \cup \ldots \cup B_i^j$ are empty sets), where $B_i$ are base symbols occurring in $T$, then $I$ is called a trivial model. Given a T-Box $T$ with terminological cycles, the problems that we are interested in are as follows:

1. Does $T$ have a nontrivial model?
2. How to construct a nontrivial model or a class of models?
3. What is the complexity of subsumption reasoning algorithm between defined concepts $A$ and $B$, where $A$ and $B$ are defined symbols occurring in T-Box $T$?

At present, the best result of problem (1) is the following: if $T$ is negation free terminology, then $T$ has the greatest fixpoint models (denoted by gfp-models) and the least fixpoint models (denoted by lfp-models), respectively. However, how to extend this result to the situation that the definition formulae contain negation symbols is still an interesting open problem.

With regard to problems (2) and (3), because $\Delta$ is finite, algorithms always exist. Generally, the algorithms are exponential and are not applicable. Therefore, many people have been searching for smaller systems of which the complexity of computation and reasoning are polynomial. In 1990, Nebel et al.\cite{9} noticed the importance of using automata theory to characterize the terminological cycles and studied terminological cycles by using automata and its graph representation. Nebel studied the $\mathcal{FL}_0$ system which includes two constructors: $\sqcup$ and $\forall R.C$. He has proved the following: the subsumption problem of $\mathcal{FL}_0$ without terminological cycles is coNP-hard. In 1996, Baader\cite{7} did further research on $\mathcal{FL}_0$. By using description graph $GT$ and simulation between description graphs, he has proved the following: the subsumption problem of $\mathcal{FL}_0$ with terminological cycles is PSPACE. In 2003, Baader\cite{10} studied $\varepsilon L$ system which includes two constructors, $\sqcap$ and $\exists R.C.$, and proved the following: the subsumption problem of $\varepsilon L$ with terminological cycles is polynomial.

Notice that both $\mathcal{FL}_0$ and $\varepsilon L$ are very small systems; they do not contain negation constructor. This shows that to find a nice theoretical framework of terminological cycles in general is quite difficult. On the other hand, this just means it is significant to study terminological cycles. Our goal is the following: start new directions at the primary stage, obtain some results, and then search for the extensions which include the works of Nebel and Baader et al.

We will define a new description logic $\nu L$ and another version of the description graphs $GT$ and $GJ$. A syntax condition for the satisfiability of membership relation will be given. By using this syntax condition, we prove the following: The subsumption reasoning in $\nu L$ with respect to