An UML Model Query Method Based on Structure Pattern Matching

Xuelin Zhang, Huajie Chen, and Tian Zhang

State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
Department of Computer Science and Technology, Nanjing University, Nanjing, China
{zhangxl,chenhj}@seg.nju.edu.cn, ztluck@nju.edu.cn

Abstract. UML has been widely used for modeling, and models are becoming increasingly important in the software development process. As a consequence, the number of models being used is increasing. Comprehending and reusing models face a real challenge. So, it is necessary to get efficient methods to query models. In this paper, we propose an UML model query method which is based on structure pattern matching. This paper will show how to extract the structure information of UML model from the textual model file. Then, a model query language and the matching algorithm will be detailed described. After that a case study is presented, which proves the effectiveness of our query method.

Keywords: model query, structure pattern matching, query language.

1   Introduction

UML is the OMG's standard for object oriented modeling and has quickly become the de facto standard for specifying OO systems[1]. UML has been widely used for modeling at different phases of a system. With model-driven development (MDD) gaining importance, an increasingly large number of models are being produced and used by software organizations. The model repository becomes so large that it is too difficult to be comprehended. Moreover, the reuse problem is extended to models in addition to code and other artifacts[2], because of that the difficulties in finding reusable models are greater. So, it is necessary to get efficient methods to query models in large repositories. Querying models can not only help to comprehend and reuse models of the system but also have an important educational benefit for students and teachers, allowing learners to find good or bad examples that can then be used to improve the acquisition of knowledge and provide hints at solutions[2].

UML sustains many aspects of software engineering. However, it does not provide an explicit facility for writing queries[1][10]. The Object Constraint Language (OCL) is a precise text language that provides constraints and object query expressions on UML model that cannot otherwise be expressed by diagrammatic notation[9]. Since the specification of the OCL mentions "query language" as one of its possible applications, lots of researchers try to form a query language based on OCL and

extensions to it[3][4][5][6][7]. Querying models with OCL query language, which will generate very complex query statements[5] and with poor efficiency, is not a good choice. The developers must be skilled in OCL programming before they can query models with OCL language. What's more, OCL was originally designed specifically for expressing constraints about a UML model, it is not a specialized language for querying[7][9]. Hence, the query expressions formed with OCL even don't have the same expression power as those formed with relational algebra[1].

The OMG defines a standard document type definition (DTD)[11] for the UML model file. So, the UML model can be described in a XMI type document, which follows the DTD standard. In this way, many tools of the UML can exchange data easily with the common DTD standard. Because of this, we can get the model information from the exported XMI model file. Then, some query work can be done with the model information without OCL. In this paper, we proposed a method of UML model query, which is based on structure pattern matching. A target model pattern and the system UML model are needed before querying. Then, after parsing the target model, it will be matched with the system UML model using a specialized algorithm. In addition, the target model pattern is described with a query language in a textual file, which is defined by us. The framework of model query is showed as figure 1.
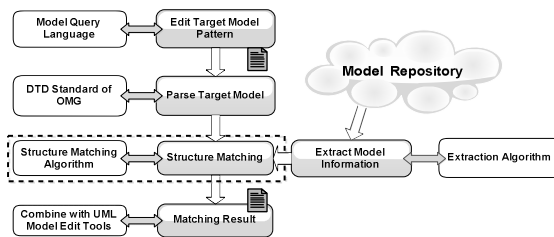

Fig. 1. Framework of Model Query

The rest of this paper is organized as follows: Section 2 discusses the method of extracting the structure information of UML model. In section 3 a model query language will be described. Section 4 discusses the algorithm of model query and its implementation. Section 5 shows a case study of our model query. Section 6 concludes the paper and presents our plans for future work.

2   Extraction of Model Information

The system UML model is described in a XMI type document. However, if you attempt to read the model file line by line, you will find that it isn't practical to comprehend the whole structure of the system model because of its complex document structure. The model file consists of very abundant contents, which contain lots of information that we don't need. As the model file is a XML type document follows the DTD standard, we chose a XML parser that based on SAX[12] to parse