An Empirical Study of the Nesting Level of Composite States Within UML Statechart Diagrams

José A. Cruz-Lemus[1], Marcela Genero[1], Mario Piattini[1], and Ambrosio Toval[2]

[1] ALARCOS Research Group,
Department of Computer Science, University of Castilla – La Mancha,
Paseo de la Universidad, 4 – 13071 Ciudad Real, Spain
{JoseAntonio.Cruz, Marcela.Genero, Mario.Piattini}@uclm.es
[2] Software Engineering Research Group,
Department of Computer Science and Systems, University of Murcia,
Campus de Espinardo – 30071 Murcia, Spain
atoval@um.es

Abstract. As UML statechart diagrams are the core for modeling the dynamic aspects of software systems, we have been studying their understandability for the last three years. In previous researches, we have already studied the relationship between many of the constructs of the UML statechart diagrams and the effect that they have on the understandability of the diagrams themselves. We have also performed a family of experiments whose results indicated that the use of composite states make UML statechart diagrams easier to understand. This fact motivated us to go a step further and investigate if the Nesting Level of Composites States (NLCS) has an impact on the understanding of the diagrams through a controlled experiment and a replication. In this paper, we present the experimental process and the main findings of them. Unfortunately, the obtained results are not quite conclusive and we have not been able to find an optimal use of nesting within UML statechart diagrams and further empirical research is needed, considering more complex UML statechart diagrams.

1 Introduction

New approaches in software engineering like MDA (Model Driven Architecture) [17] and MDD (Model Driven Development) [1] are enabling a shift in focus from software to models of software. These approaches consider models as end-products rather than just mean to produce software.

In truly 'model-driven' software engineering, the quality of the models used is greatly important. For that reason, models like UML ones are gaining more relevance in the development of software, as the quality of the models used will later determine the quality of the software systems produced.

As UML statechart diagrams are the core for modeling the dynamic aspects of software systems [13], we have been studying their understandability for the last three years. Our main idea was that if diagrams are difficult to understand this will affect their maintainability. In previous researches, we have studied the relationship between many of the constructs of the UML statechart diagrams and the effect that they have

J. Akoka et al. (Eds.): ER Workshops 2005, LNCS 3770, pp. 12 – 22, 2005.

on the understandability of the diagrams themselves. First, we defined and validated a set of metrics [11] for evaluating if the structural properties of UML statechart diagrams, such as size and complexity, influenced the understanding of UML statechart diagrams. In these researches we had found that the usage of composites states had apparently no influence on the understandability of UML statechart diagrams. This fact seemed to be a bit suspicious. For that reason, we decided to run another experiment, and a further replication, for specifically studying if the use of composite states facilitated or not the understanding of UML statechart diagrams [10]. The results of this empirical study indicated that the use of composite states improves the understandability efficiency of UML statechart diagrams, i.e. how accurately the different stakeholders understands the diagrams, if the subjects have a certain level of experience in working with this kind of UML diagrams. These findings motivated us to go a step further and define a new metric named *Nesting Level in Composite States* (NLCS) which indicates the maximum number of nested composite states in an UML statechart diagram. We based on the measure DIT (Depth of Inheritance Tree) defined in [9], as we think that there is a certain similarity between the NLCS within an UML statechart diagram and the depth of a within a generalization hierarchy in an UML class diagram.

In this paper, we will investigate in the NLCS affects the understanding of UML statechart diagram and try to find the optimal nesting level within a diagram through a controlled experiment and a replication of it. For designing the experiment we took several ideas from the different experimental experiences performed related to the DIT metric [3-8, 12, 14, 18, 19, 21].

Not only do we want this paper to be taken under a research point of view, but also to be useful for designers and software engineering teachers at universities.

We will begin defining our research question. The description of the experimental process, covering the design, tasks and performance of the experiment is explained in section 3. Section 4 describes the data analysis and the interpretation of the obtained results. Section 5 tackles all the features related to the replication of the experiment. Finally, conclusions and future work are presented in section 6.

2 Research Question

As we commented in section 1, our research question can be stated as:

> *Does the use of different nesting levels of composite states in UML statechart diagrams affect the understandability of the diagrams?*

In order to answer this question we have defined the previously presented metric NLCS. Based on the guidelines exposed in [18], we have formulated the following experimental hypotheses:

- $H_{0\text{-}ij}$: the understandability of UML statechart diagrams with i and j composite states nesting levels is not significantly different,
- $H_{1\text{-}ij}$: the understandability of UML statechart diagrams with i and j composite states nesting levels is significantly different,

  In both cases, $i, j \in \{0, 1, 2\}$ and $i \neq j$.