INTRODUCTION

According to Booch, Rumbaugh, and Jacobson (1999), developing a model for an industrial strength software system before its construction is regarded increasingly as a necessary activity in information systems development. The use of object-oriented modeling in analysis and design started to become popular in the late eighties, producing a large number of different languages and approaches. Lately, UML (2004) has taken a leading position in this area.

In this article, we give an overview assessment of UML using a generic evaluation framework. We will first present the evaluation framework. We will then evaluate the language quality of UML before pointing to the future direction and potential of UML.

BACKGROUND

Krogstie, Sindre and Lindland (1995) and Krogstie and Sølvberg (2003) have developed a framework for quality of models and modeling languages.

The main concepts of the framework and their relationships are shown in Figure 1 and are explained in the following. Quality has been defined referring to the correspondence between statements belonging to the following sets:

- G, the goals of the modeling task.
- L, the language extension, that is, the set of all statements that are possible to make according to the graphemes, vocabulary, and syntax of the modeling languages used.
- D, the domain, that is, the set of all statements that can be stated about the situation at hand.
- M, the externalized model itself.
- K_M, the relevant explicit knowledge of those being involved in modeling. A subset of these is actively involved in modeling, and their knowledge is indicated by K_M.
- I, the social actor interpretation, that is, the set of all statements that the audience thinks that an externalized model consists of.
- T, the technical actor interpretation, that is, the statements in the model as “interpreted” by modeling tools.

Figure 1. Framework for discussing the quality of models
The main quality types are indicated by solid lines between the sets, and are described briefly in the following:

- **Physical quality**: The basic quality goals on the physical level are externalization, that the knowledge $K$ of the domain $D$ has been externalized, and internalizeability, that the externalized model $M$ is available.

- **Empirical quality** deals with predictable error frequencies when a model is read or written by different users, coding (e.g., shapes of boxes) and HCI-ergonomics for documentation and modeling-tools. For instance, graph layout to avoid crossing lines in a model is a mean to address the empirical quality of a model.

- **Syntactic quality** is the correspondence between the model $M$ and the language extension $L$.

- **Semantic quality** is the correspondence between the model $M$ and the domain $D$. This includes validity and completeness.

- **Perceived semantic quality** is the similar correspondence between the audience interpretation $I$ of a model $M$ and his or hers current knowledge $K$ of the domain $D$.

- **Pragmatic quality** is the correspondence between the model $M$ and the audience’s interpretation of it ($I$). We differentiate between social pragmatic quality (to what extent people understand the models) and technical pragmatic quality (to what extent tools can be made that interpret the models).

- The goal defined for social quality is agreement among audience members’ interpretations $I$.

- The organizational quality of the model relates to that all statements in the model contribute to fulfilling the goals of modeling (organizational goal validity), and that all the goals of modeling are addressed through the model (organizational goal completeness).

Language quality relates the modeling languages used to the other sets. Four quality areas for language quality are identified, with aspects related both to the meta-model and the notation as illustrated in Figure 2.

1. Domain appropriateness. This relates to the language and the domain. Ideally, the conceptual basis must be powerful enough to express anything in the domain, not having what (Wand & Weber, 1993) terms construct deficit. On the other hand, you should not be able to express things that are not in the domain, that is, what is termed construct excess (Wand & Weber, 1993). Domain appropriateness is primarily a mean to achieve physical quality, and through this, to achieve semantic quality.

2. Participant language knowledge appropriateness relates the participant knowledge to the language. Participant language knowledge appropriateness is primarily a mean to achieve physical and pragmatic quality.

3. Participant comprehensibility appropriateness relates the language to the social actor interpretation. The goal is that the participants in the modeling effort using the language understand all of the possible statements of the language. Comprehensi-