IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Hershey PA 17033-1117, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB8080

Chapter XIV

A Unifying Translation of Natural Language Patterns to Object and Process Modeling

Alexandra Galatescu
National Institute for R&D in Informatics, Romania

ABSTRACT

The proposed translation of natural language (NL) patterns to object and process modeling is seen as an alternative to the symbolic notations, textual languages or classical semantic networks, the main representation tools today. Its necessity is motivated by the universality, unifying abilities, natural extensibility, logic and reusability of NL. The translation relies on a formalized, stylized and graphical representation of NL, bridging NL to an integrated view on the object and process modeling. Only the morphological and syntactic knowledge in NL is subject to translation, but the proposed solution anticipates the semantic and logical interpretation of a model. A brief presentation and exemplification of NL patterns in consideration precede the translation.

INTRODUCTION AND MOTIVATION

The complexity of today's information systems (IS) entails the need to combine and integrate object, process, data flow, workflow models and, at the same time, to satisfy *modeling requirements* like: universality, integration, extensibility, equidistance versus the implementation models, structuring (modularity, connectivity, encapsulation), logical and semantic consistency, reusability, friendly interface. Symbolic notation has become the main representation tool for most of these models

(Harmon, 1997; Johannesson, 1996; Lawrence, 1997; Lee, 1997; Taylor, 1995; OMG, 1997a; 1997b). But, it has important *disadvantages* regarding:

- *model extensibility*, because any new type of operators, concepts, relationships, etc., introduces a new symbol, icon, or marker, resulting into an overwhelming notation, e.g. Unified Modeling Language (UML) (OMG, 1997a; 1997b);
- *model learning*, because too many distinct symbols increase the designers' confusion;
- *system internal structure and evolution,* because there are not universal rules to uniformly translate symbols with heterogeneous semantics and logic into internal structures, irrespective of the implementation models;
- *control of the system logical and semantic consistency* during early phases of the system's life cycle, because there are not universal rules to generalize and formalize, hence to control, the logic and semantics of the correlated symbols;
- *integration of the object and process models* because, usually, there is not a sound and generalized solution to the seamless integration of the symbols from the semantic and logical points of view (see UML).

An alternative solution and a step toward natural language-oriented (NL-oriented) modeling are the *textual languages* (e.g. the formal language proposed for KADS (Schreiber, 1993) or the language in Moral, (1998)), necessarily accompanied by textual parsers. Unfortunately, these languages either lack universal rules to formalize, control and integrate the semantics and logic of the models, or use complicated and specific formulae/ phrases (usually, strings of symbols) to describe them.

Instead, this paper examines the patterns of our primary, universal and vital abstraction of real life, natural language and tries to synthesize, stylize, adapt and apply these patterns to IS modeling. The universality, as well as the morphological and syntactic stability of a NL-oriented model, facilitate the *communication* among distributed programs, IS, users, Internet sites, etc. (Lewrenz, 1999; Steuten, 1999; Thalheim, 1999) and are the solution to the modeling requirements enumerated above.

A next step (beyond the scope of the paper) is to analyse, then to synthesize, adapt and apply the most important and general ideas, rules, algorithms, etc., already discussed and accepted in NL processing theories and tools and appropriate to object and process modeling.

A formalized, stylized and graphical representation of NL, intermediating the conceptual transfer of NL patterns to object and process modeling, is the main result of the paper. This transfer is seen here as part of the broader correspondence between the types of knowledge in NL and an extended ontological and epistemological background of the complex IS.

Related research. Many theoretical or practical results have been recently obtained with respect to the application of NL paradigm to IS engineering. The most important research directions are: the linguistic interpretation of the models (mainly ER, e.g., Lewrenz, 1999), the modeling of the systems' dynamics (e.g., Burg, 1995; Steuten, 1999), the human-computer interaction (e.g., Lewrenz, 1999), the requirements engineering (e.g., Fliedl, 1999), the organization modeling (e.g., Medina-Mora, 1992; Steuten, 1999), the knowledge representation (e.g., Sowa, 1984; 1988;

From Knowledge to Personal Knowledge Management

Fortunato Sorrentino (2010). *Information Resources Management: Concepts, Methodologies, Tools and Applications (pp. 157-165).*
www.irma-international.org/chapter/knowledge-personal-knowledge-management/54477