Specifying Participants Behaviour in Generalized Eventity Frames

MILENA SLAVCHEVA
(IICT, Bulgarian Academy of Sciences)

This paper presents the construction of an application-driven and task-motivated taxonomy of STATES necessary for specifying the semantic descriptors of verbal units. The concept hierarchy is used for building semantic descriptors, which populate *SemInVeSt* (*Sem*antically *In*terpreted *Ve*rb-centred *St*ructures) - a knowledge base of the semantics of verbs (Slavcheva, 2008). So far a reflexive-verb-component of *SemInVeSt* has been built, containing verbs in a reflexive form in Bulgarian and their semantic equivalents in French and in Hungarian. The verbs are provided with EVENTITY FRAME TEMPLATE diagrams (Figure 1), where the characteristic features of the eventity participants and their behaviour are represented. The EVENTITY FRAMES in *SemInVeSt* are constructed in conformance with the *Unified Eventity Representation* (UER) (Schalley, 2004) - a cognitive theoretical approach to verbal semantics and a graphical formalism, based on the *Unified Modeling Language* (UML) (OMG 2005). The focus in this paper is on the dynamic core of the EVENTITY FRAMES (denoted by the dashed-outline rectangle with rounded corners in Figure 1), which contains the STATE-machines of the prominent participants and models their behaviour and interaction. At present the EVENTITY FRAMES, associated with the verbal units, are TEMPLATES - they contain parameterized STATES (indicated by *A* in the right-hand side STATE-machine and the small dashed-outline rectangle in the upper right part of the EVENTITY FRAME octagon). That means that the semantic descriptors are generic ones. In order to construct EVENTITY FRAMES specified for each verbal unit in the knowledge base, the parameters have to be bound, that is, the STATES have to be given names and assigned properties.

The UER metamodel inheritance structure provides a level of basic eventity types, whose basic building blocks are the SIMPLE STATES and the TRANSITIONS. The concept of STATE is defined as "a condition during the life of an object or an interaction during which it satisfies some condition, performs some action, or waits for some event" (OMG 2005). The change of the state of a given object is represented by a TRANSITION. The SIMPLE STATES are subdivided into: 1) PASSIVE SIMPLE STATES (PSS) in which the participant satisfies some condition and is characterized as being passive; 2) ACTIVE SIMPLE STATES (ASS) in which the participant performs some action and is characterized as being active. The ACTIVE SIMPLE STATES can further be subdivided into: 1) ACTS in which the action is non-durative, punctiform; 2) ACTIVITIES in which the action is considered ongoing, durative. (Schalley, 2004) Thus the classification of the basic eventity types in UER differs from that of (Vendler, 1957) or (Pustejovsky, 1995).

The specification of the STATES depends on the modeling granularity determined for a given representation. The basic means for specifying the semantic descriptors are the following.

1. A set of STATE names is defined.
2. Clusters of PROPERTIES are determined, which further specify the STATES where necessary.

3. STEREOTYPES and *keywords* are defined, used for the formulation of subsets of modeling elements where necessary.

Here arises the problem of the so called semantic primitives. Since it is difficult (and widely considered as unnecessary) to define a very small finite set of (possibly indecomposable) elementary senses (cf. Apresjan, 1974, 2000; Wierzbicka, 1996; Goddard 2002; Gladkova, 2007; Schalley, 2004), I take a "looser approach" in choosing basic concepts for the decompositional semantic descriptions. Thus I use as a supplier of basic concepts a *Semantic Dictionary - Minimum* (Kasabov, 1990), which claims to provide the semantic nucleus of the lexical system in Bulgarian. I propose an initial, first, version of taxonomy of basic concepts used for specifying the EVENTITY FRAME TEMPLATES of the reflexive-verb-component of *SemInVeSt*. The taxonomy construction adheres to the UER representational framework, which allows the user to plug in description elements according to his needs. The UER description inventory is extended as follows.

1. The set of named basic STATES is entirely user-defined. The PASSIVE SIMPLE STATES are instantiated with instances like *Be, Feel, Know, Slow, Fast, High, Low, Big, Small, Long, Short, Dirty, Clean, Young, Old, New, Warm, Cold, Wide, Full, Empty, Narrow, Wet, Dry, Ready*, etc. The ACTIVE SIMPLE STATES are instantiated with instances like *Do, Make, Become, Perceive, Think, Talk, Move, Keep*, etc.

2. The PROPERTIES, which are part of the metamodel, further specify the STATES. The values of the PROPERTIES are usually of the ENUMERATION or *Boolean* data type. The predefined PROPERTIES of UER, which apply to the STATES, are: *default* (type Boolean), *manner* (type ENUMERATION: leaping, etc.) and *sensoryOrgan* (type ENUMERATION: eye, ear, etc.). A number of user-defined PROPERTIES have been added to the metamodel like *partBody* (type ENUMERATION: neck, knee, etc.), *categoryEthics* (type ENUMERATION: guilt, etc.), *emotion* (type ENUMERATION: anger, joy, etc.), *direction* (type ENUMERATION: up, down, back, etc.), etc.

3. The predefined UER STEREOTYPES, which apply to the STATES, are: <<repetitive>>, <<be-at>>, <<move-along>>, <<aggregated>>. STEREOTYPES specifying location relations are added like <<be-inside>>, <<be-outside>>, <<be-near>>, <<be-far>>, etc.

For the sake of being more objective in constructing the conceptually grounded semantic descriptors, the taxonomy of basic concepts, as well as the eventity descriptors, are compared with widely popular ontologies like *OpenCyc*[1] and SUMO[2] (Niles and Pease, 2001). It should be noted that the comparison between *SemInVeSt* and the ontologies, as well as the borrowing of some concepts from the ontologies, is biased in favour of the linguistic nature and UML-based representation of *SemInVeSt* as opposed to the encyclopedic character and specific knowledge formalization of the ontologies.

Apresjan, J. (1974). Lexical semantics. Moscow: Nauka.

Apresjan, J. (2000). Systematic lexicography. New York: Oxford University Press.

Gladkova, A. (2007). Universal and language-specific aspects of propositional attitutes: Russian vs. English. In: Schalley, A. & Khlentzos, D. (eds.). Mental states. Volume 2: Language and cognitive structure. Amsterdam / Philadelphia: John Benjamins, 61-63.

[1] http://opencyc.org/

[2] http://www.ontologyportal.org/

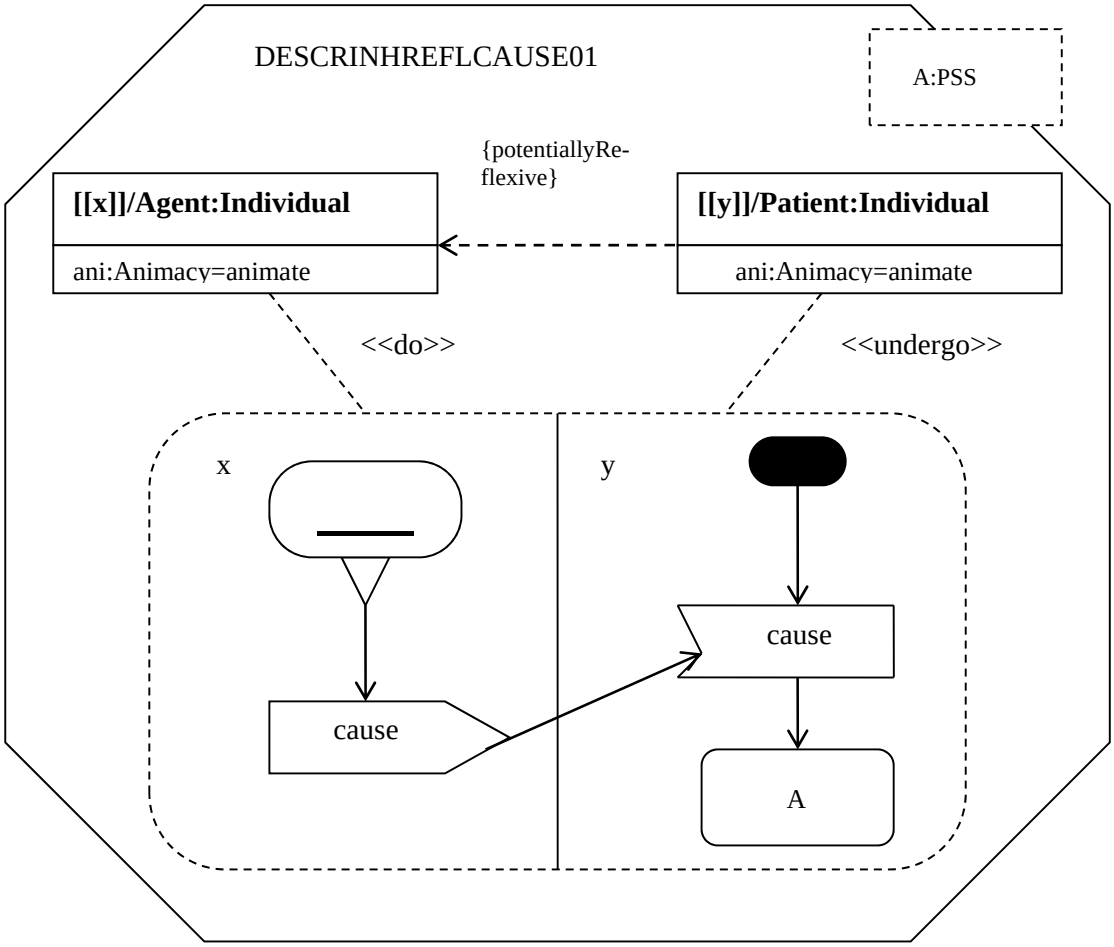

Figure 1: Generalized EVENTITY FRAME representing a sub-class of the inherent-reflexive-verbs class

Goddard, C. (2002). The search for the shared semantic core of all languages. In: Goddard, C. & Wierzbicka, A. (eds.). Meaning and universal grammar – Theory and empirical findings, Volume I. Amsterdam: John Benjamins, 5-40.

Kasabov, I. (1990). Semantic dictionary – minimum. Sofia: Kliment Ohridski University Publishing House.

Niles, I. & Pease, A. (2001). Towards a Standard Upper Ontology. In: Welty, C. & Smith, B. (eds.). Proceedings of the 2nd international conference on Formal Ontology in Information Systems (FOIS-2001). Maine: Ogunquit.

OMG. (2005). ISO/IEC 19501: Unified Modeling Language Specification, Version 1.4.2. Object Management Group (OMG). http://www.omg.org/.

Pustejovsky, J. (1995). The generative lexicon. Cambridge / London: The MIT Press.

Schalley. A. (2004). Cognitive modeling and verbal semantics. A representational framework based on UML. Berlin / New York: Mouton de Gruyter.

Slavcheva, M. (2008). Thinking in objects: Towards an infrastructure for semantic representation of verb-centred structures. Lecture Notes in Artificial Intelligence, Volume 5246. Berlin / Heidelberg: Springer, 193-200.

Vendler, Z. (1957). Verbs and times. The Philosophical Review, Vol. 66, No. 2, 143-160.

Wierzbicka, A. (1996). Semantics: primes and universals. New York: Oxford University Press.