ONTOLOGICAL FEATURES OF ENTITIES IN MOTION EVENTS AND THEIR ROLE IN THE SEMANTIC ANALYSIS OF SENTENCE

European Union
European Social Fund
Investing in your future
EXCS ESTONIAN COMPUTING

Haldur ÕIM, University of Tartu, Estonia, haldur.oim@ut.ee

UNIVERSITY OF TARTU 1632

GENERAL FRAMEWORK AND AIMS

The linguistic-semantic descriptions of language expressions do not provide results which satisfy the needs of language technology applications. One kind of additional information needed is **ontological knowledge of the domain (shared knowledge, common ground)**. Our domain is physical motion: X is moving from place P1 to Place P2, agentively, non-agentively: ***walking, flying, swimming, putting, pushing, throwing, rolling, sliding***…

The central concept here is not a sentence (or text) but ***Event***, the relevant components of which can be 'collected' from the text, ***or*** should be supplied by the ontological representation of the Event in a database. We use a variant of frame semantics, where the central unit of an ***Event frame*** is ***Predicate***, and the ***participants and situational aspects involved are Arguments in definite Semantic Roles***.

Our main aim here is, ***first***, to show that ontologically most *Events*, linguistically described as coherent wholes, ontologically constitute complexes of *Subevents*. And **second**, to offer a treatment of the entities that **participate** in these event/subevents. These entities are divided in the *Motion Participants*, entities that actually participate in the motion, and the *Motion Space* - the environment where the motion takes place. In the case of motion events, the ontological knowledge often comes into the picture just through these last entities: who is walking where, who is throwing what, whereto, etc).

EVENTS AND SUBEVENTS: AN EXAMPLE

Example sentence: John threw a stone from the road into the bushes.

A **throwing-event** is usually described as: an *Agent* (John) causes [by the movement of his hand (*Instrument*)] an *Object* (stone) to move from P1, *Locfrom* (the road) to P2, *Locto* (the bushes) through the air (*Path*). The roles *Instrument* and *Path* represent here, according to linguistic semantics, so-called hidden arguments not present in the sentence. But even adding them to the frame does not make the structure of the event explicit: it is an **Event** constituting a series of **Subevents** which are connected with each other by their **Consequence - Prerequisite** components (below we show only three of them, E1 - E3, although they in fact are complex, too).

E1: A(gent) is Picking up O(bject) from P1 by his hand (> A is Holding O in his hand);

E2: A Causes, by Moving his hand, O to Leave the hand and Move through the air (> O is Moving through the air);

E3: O Falls down to P2.

The next crucial problem is that depending on the fillers of the roles A, O, P1, P2 the ontological contents of the whole event can be quite different: is A an grown-up, a child, or even a chimpanzee, what is O (and if it's a stone, then how large), what are P1 and P2 (falling down into bushes is different from falling down into a lake). And what about the **Distance** from P1 to P2 (centimeters, meters, etc)?

MOTION PARTICIPANTS AND MOTION SPACES

There is no clear difference, in ontology, between the entities which can function as participants (moving themselves, causing others' motion) and as environments where the motion takes place. Nevertheless, there are, **first**, entities that in our world picture function *typically and generally* as Motion Spaces (e.g a forest, a lake, a road, a river), and that as Motion Participants (e.g. living beings, stones or movable artifacts). **Second**, in case of concrete event types this division can definitely be made more clear (cf events like walking, swimming, flying or rolling).

For **Predicates** (= **Events**), frame semantics offers an elaborated conceptual system of categories and semantic roles. The same kind of system of conceptual means is needed for representing **Motion Participants and Motion Spaces**. In cognitive semantics there already exist some classifications, e.g Talmy's distinction between **Figure** and **Ground**. It is this approach we have chosen as the conceptual basis for our project. And as the framework for formal representation we have chosen the **qualiastructure** approach (Pustejovsky a. o.), the reason being that its representation form can easily be incorporated into the predicate-argument/role structures of frames. BUT the known qualia-system (***Formal, Costitutive, Telic, Agentive***) should be elaborated considerably in order to be used in the description of ***Motion Spaces and Motion Participants. Some Examples***.

MOTION SPACES (MS)

An MS can be a field, a road, a house, a kitchen, a forest, a river, a lake etc. On this basis already, we can make a difference between ***Planes*** and ***Spaces/Containers*** - two-dimensional entities ON which Participants can move, and three-dimensional entities IN/INTO/OUTOF which they can move. Both types of MS can be elaborated further, e.g in the line of the *Costitutive quale*: in a forest there are trees, in a lake there is water, a road differs from a field by the additional dimension *Linearity*, and these properties restrict possible ways of motion for Motion Participants.

MOTION PARTICIPANTS (MP)

Here it is much harder to identify typical dimensions for classifying theMPs. Thus we have chosen the way of working out a system of ***dominating*** properties on the ground of which the entities can be cross-classified: among physical objects we can differ between **Form-dominated** entities (like stones) and **Function-dominated** entities (artifacts, body-parts), etc. And **secondly**, especially in case of artifacts, it is important to differentiate between certain **States**: a glass can be full or empty, a door can be open or closed, and this can be relevant for a motion event.

A VERY SHORT SUMMARY

The description of the problems given above should have made it clear that serious applications, for instance using robots able to communicate in natural language, are possible, at present, only if the structure of the (sub)domain is very clearly delimited and its internal structure described in detail.

Reference: H. Õim, H. Orav, N. Kahusk, P. Taremaa, Semantic analysis of sentences: the Estonian experience. In: Human Language Technologies - The Baltic Perspective. Ed by I. Skadina and A. Vassiljevs. Amsterdam etc: IOS Press 2010, 208-216.