Event Representation in Text Understanding
Transfer of Meaning Structures

Haldur Õim[1] and Mare Koit[2]
[1]*Institute of Estonian and General Linguistics, University of Tartu, Jakobi 2, Tartu, Estonia*
[2]*Institute of Computer Science, University of Tartu, Liivi 2, Tartu, Estonia*

Keywords: Meaning, Frame, Event, Physical Motion, Communication.

Abstract: When modelling language understanding we have to deal with the process of transferring meanings. Humans cognize and organize the knowledge about the world, physical as well as social, in such categories as objects, situations, processes, events, etc., not sentences. The same should hold in a computational model. In this paper we will consider one kind of these categories, events. We will discuss the possible analogy in structuring the physical and social events and, accordingly, the possibility to use analogous conceptual and formal means to represent them.

1 INTRODUCTION

When we want to transfer the text understanding ability to the computer, we must transfer a theory, a model of how the understanding system works in humans because texts we are dealing with are written by humans. The problem pertains to the 'units of understanding'. Formally, texts are concatenations of sentences which are built according to certain (language-specific) rules. The upper linguistic level which deals with sentences is syntax but when modelling language understanding we must go into semantics and pragmatics which deal with the process of transferring meanings (information about the domain dealt with in the text, and the communicative intentions of the author of the text). The first point we want to stress is that here, in describing the understanding process and, in particular, its results, other units than sentences are needed. Humans cognize and organize the knowledge about the world, physical as well as social, even when they get this knowledge from texts, in such categories as objects, situations, processes, events etc., not sentences. The same should hold in a computational model. We must have formal representations of such knowledge in order to use it.

In this paper we will deal with one kind of the necessary categories, events. Events constitute a rather specific category in organizing our knowledge about the world. In a sense, we 'impose' these structures on the continuous flow of what happens around us. Typically, we remember and talk about the past in terms of events: they have definite inner structure, starting and ending states, they can be organized hierarchically, contain other events as parts (Tversky et al., 2011). Further, events are domain-specific. For instance, events of the physical and social world can be quite different in details. This means that in understanding and representing them we have to use not only linguistic, but also ontological information. On the other hand, events of different domains can have much in common. In the cognitive approach to language understanding it is a commonly accepted thesis that knowledge of abstract domains is regularly structured using the structuring of some more concrete domains.

This latter thesis constitutes one of the main background topics we want to deal with in the present paper. We have, for quite a long period, dealt with two domains: (1) motion, agentive as well as non-agentive (a physical domain) and (2) human interaction, especially dialogues (a social domain) (Õim et al., 2010; Koit et al., 2006). Here we will discuss the possible analogy in structuring the corresponding events and, accordingly, the possibility to use analogous conceptual and formal means to represent the events of both domains. Of course, it is clear that dialogues as social events have more complex and specific structure than motion events. But the general conception underlying both

Õim H. and Koit M..
Event Representation in Text Understanding - Transfer of Meaning Structures.
DOI: 10.5220/0004625703670372
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2013), pages 367-372
ISBN: 978-989-8565-81-5

of them has much in common. For instance, both are dynamic domains. In particular, in dialogues also 'something' is moved (from one participant to another). And since the domain of motion has been studied in more detail in linguistic as well as in ontological semantics, it is relevant to ask whether the results attained here could be used in dealing with dialogues.

The paper is organized as follows. In Section 2 we consider the structure and representation of events using frames and qualia structures. Section 3 investigates motion events and Section 4 communication events. Section 5 discusses some problems and in Section 6 we will make conclusions.

2 EVENTS AND FRAME SEMANTICS

The description of the general theoretical framework of our work can be found, for instance in (Õim et al., 2010, Õim, 2012). We are using a kind of frame semantics approach together with qualia structure approach provided by J. Pustejovsky (1995) and R. Jackendoff (2002). Events described in sentences are represented as frames.

The original idea behind the concept of frame came from frame semantics and specifically from FrameNet (see e.g. Fontenelle, 2003 for overview). Still, for purposes of our study we had to work out our own inventory of semantic roles. One reason for this was the need to draw inferences from frames: FrameNet does not deal with inferences, at least not explicitly. In case of semantic analysis of text it is impossible to ignore this problem; and certain kinds of inferences are directly connected with semantic roles. This, by the way, does not mean that FrameNet structures cannot be used to draw inferences from sentences. We have tried this, in parallel with our frame structures. But the role inventory in FrameNet is too complicated and domain-dependent to be taken as a regular basis of sentence/text semantic analysis program at the very beginning.

Frames in our system are structures consisting of a head – a (motion) verb which in a sentence can function as predicate – and its possible arguments as fillers of certain semantic roles. Thus, semantic roles are the main structuring elements of a frame. Although the heads of frames are verbs, the frames are in fact not frames of verbs but frames of events represented/designated by the verbs as possible predicates of corresponding sentences. The basic semantic unit in text semantics is not a word, nor even a sentence, but an event (in our domain of motion). The details of one such event can be picked up from different sentences, they should be collected and integrated into the frame of this individual event. A frame is a semantic description of a predicate, including all of its possible arguments and their semantic roles. From the point of view of semantics, the arguments represent participants of an event and the predicate determines type and the general structure of the event. Each participant (argument) has a certain role in the structure.

We are considering complex events described in texts – the events which include sub-events related to each other in a certain way, e.g. temporally or causally. Complex events express dynamics of texts – which sub-events will cause other sub-events, which previous sub-events have to be occurred before a certain sub-event, etc.

Our research data have thus far come mainly from Estonian but we are compared them with data from other languages (English, Russian, a. o.).

Related work goes back at least to seventies, e.g. to the work done by Roger Schank and others on motion and communication in the frames of modelling 'story understanding' (Schank, 1975, 1986). His conceptual dependency theory states that all conceptualizations can be represented in terms of a small number of primitive acts performed by an actor on an object (e.g. MTRANS for transferring mental objects and ATRANS for transferring physical objects). Events are understood in terms of scripts, plans and other knowledge structures as well as relevant previous experiences. (Shank and Abelson, 1977). Since eighties, the topic of meaning transfer has intensively been studied in the conceptual metaphor theory (e. g. Lakoff, 1987).

In the next section we will give an overview of our conceptualization of motion events. Since we are interested in using motion as a 'source domain' for structuring (human) communication as the 'target domain', not all details of motion events are of equal importance. For instance, several physical characteristics of the entities participating in a motion event are not relevant in case of communication events and are omitted here.

3 MOTION EVENTS

The critical difference between an event taken 'as a whole' and the 'pure act' of motion as denoted by the corresponding predicate or an isolated sentence

lies in the need to take into account also the contextual information which should hold and must be made explicit.

Thus, the meaning of the verb (predicate) 'to throw' as used e.g. in the sentence *John threw a stone from the road into the bushes* is usually described as: an Agent (John) caused an Object (stone) to move from one place, Locfrom (the road) to another place, Locto (the bushes). In more detailed definitions it is added: through the air (Path). Instrument (the hand of the Agent) and Path (the air) represent so-called hidden arguments (see Jackendoff, 2002, Õim, 2012). But even adding them to the throwing-event frame does not make the whole event as expressed in the above sentence ontologically explicit because in fact we have here a complex event, i.e. an event containing sub-events some of which may, in the general-semantic sense, not constitute an obligatory part of throwing. For instance, before the throwing act proper John must have picked up the stone from the road, etc.

```
FRAME: AGENTIVE_MOTION
HYPERONYM: MOTION
PRECONDITIONS: Agent has Object
ACT:
ROLE STRUCTURE
Participant Roles
ROLE: Agent (the instigator of the
event)
ROLE: Object (the entity which
moves)
          FRAME: Location_1
          Object = Object
          Loc = Locfrom
          Time = Timefrom

          FRAME: Location_2
          Object = Object
          Loc = Locto
          Time = Timeto
ROLE: Instrument (e.g. with hands)
ROLE: Locfrom (place where the
motion starts)
ROLE: Loc (place where the motion
occurs)
ROLE: Locto (place where the motion
ends)
ROLE: Direction
ROLE: Path (e.g. over a bush)
ROLE: Manner (e.g. slowly, angrily)
ROLE: Quant (e.g. how many times)
ROLE: Goal (of the Agent: Object is
located at Locto).
CONSEQUENCES: Agent does not have
Object
```

Figure 1: The frame AGENTIVE MOTION.

In Fig.1, the frame structure of AGENTIVE MOTION is given in basic details (where Agent intentionally moves Object from one place to another, as represented by verbs like 'to throw' or 'to lift'). Each role puts requirements to its fillers, e.g. Object of agentive motion has to be a physical object. There are different requirements in the case of the roles of the frames of different verbs.

The Location sub-frames are attached to the roles whose fillers move in the event described by the frame. In the described type of the motion event Object represents the only entity that obligatorily moves. Location_1 and Location_2 fix the location of the entity before and after the motion event, accordingly, taking the corresponding information from the Locfrom and Locto roles of the main frame.

In our general model, we make a distinction between *motion participants* and *motion space*, that is entities that move and that represent the 'environment' of motion. Typical motion participants are the fillers of the roles Agent, Object and Instrument, and motion space is determined by the fillers of Loc, Locfrom, Locto, Path. This is an ontological distinction specific to physical motion and, as said above, we will not consider it more closely here (but see Õim, 2012). We will return to the problem in Section 5.

In our research we are proceeding in the following way. We have chosen a number of predicates (verbs) that represent certain types of motion events (moving on the ground, in the air or in the water, using certain instruments, e.g. body-parts or vehicles, etc.). Departing from these predicates, we have collected examples from corpora and (multilingual) dictionaries, seeking for ontologically representative types of entities that can function as the fillers of the semantic roles in the corresponding motion events. The aim is to build a typology of entities that function as motion participants and motion spaces, and on this basis, the typology of motion events.

4 COMMUNICATION EVENTS

When communicating, the speakers can perform actions while making utterances. Such actions are called speech acts (asserting, commanding, requesting, etc.). The participants express certain attitudes, and the type of speech act being performed corresponds to the type of attitude being expressed. For example, a statement expresses a belief, and a request expresses a desire.

A speech act, or a communicative act, is a minimal functional unit in human communication. Every act predicts in certain degree, which another act can follow (e.g. a question has to be answered by a communication partner, a request granted, etc.). Every act can be considered as a (motion) event in communication and the dialogue itself is a complex event which includes communicative acts related to each other.

We consider communication between two participants, A and B, where the goal of A is to get some information. Our empirical material is a dialogue corpus which contains different types of dialogues, among them authentic telephone conversations (see the following excerpt of a transcribed directory inquiry taken from the corpus).

```
A: öelge palun linnaliini bussijaama
infotelefoni number.                    REQUEST
please tell me the phone number of the
town bus station
B: kolm kuus kaks     GIVING INFORMATION
three six two
A: jah   CONTINUER
yes
B: seitse kuus seitse.
                      GIVING INFORMATION
seven six seven
```

Communicative acts as motion events (used to forward and receive information) can be represented as frames where the moving object is information. The frame of the communicative act REQUEST is shown in Fig.2.

The author of the act, A (Agent) is forwarding his/her request to the addressee, B (Recipient), information (A's request) is moving from A to B. The author and the addressee themselves do not move, they can be even on different places and communicate by telephone.

Unlike of agentive motion (Fig.1), every communicative act obligatorily has two 'intentional' participants, although in different roles (Agent and Recipient). The moving object is non-physical (information), and Agent does not lose information forwarded to Recipient (which is different when moving physical objects).

Communication as a complex event can be considered as a temporal sequence of sub-events – communicative acts – and can be represented as a motion frame which contains other motion frames inside itself. In the simplest case, a dialogue consists of two communicative acts, e.g. question – answer: A asks a question and B answers it (Fig. 3).

```
FRAME: REQUEST_for_information
HYPERONYM: COMMUNICATIVE_ACT

PRECONDITIONS
        A believes that there exists d
∈ D such that p(d/x) is true
        A wants to know the element of
D which satisfies p
        A believes that B knows the
element of D which satisfies p

GOAL: B knows that A wants to know the
element of D which satisfies p

ACT: A informs B that A wants to know
which element of D satisfies p

ROLE STRUCTURE
Participant Roles
ROLE: Agent (A)
        FRAME: Location_1
        Object: Agent
        Loc: Locfrom
ROLE: Recipient (B)
        FRAME: Location_2
        Object: Recipient
        Loc: LocTo
ROLE: Object (information which moves
from A to B)
        FRAME: Location_1
        Object = Object
        Loc = Locfrom
        Time = Timefrom

        FRAME: Location_2
        Object = Object
        Loc = Locto
        Time = Timeto

ROLE: Instrument (voice)
CONSEQUENCE: B knows that A wants to
know the element of D which satisfies p
```

Figure 2: Communicative act REQUEST for information.

Information is moving from one participant to another: from A to B (question) and from B to A (answer). In the same time, both A and B keep information which has been forwarded to the partner, therefore, their knowledge is increasing in the communication process. The fillers of the roles of Agent and Recipient are changing during communication (while A and B are turn-taking).

Still, miscommunications can occur when exchanging information, e.g. Recipient does not hear or does not understand information (request) forwarded by Agent, or s/he does not have information requested by Agent (i.e. some of preconditions does not hold).

```
FRAME: EXCHANGE information
HYPERONYM: COMMUNICATION
PRECONDITIONS:
  A has INFO_A
  B has INFO_B
ACT:
  FRAME: COMMUNICATIVE_ACT_A
    AGENT: A
    RECIPIENT: B
    OBJECT: INFO_A
    INSTRUMENT: voice A
    TIME: T1
    GOAL: B has INFO_A
  FRAME: COMMUNICATIVE_ACT_B
    AGENT: B
    RECIPIENT: A
    OBJECT: INFO_B
    INSTRUMENT: voice B
    TIME: T2 (> T1)
    GOAL: A has INFO_B
CONSEQUENCE:
  B has INFO_A
  A has INFO_B
```

Figure 3: Exchange of information as a complex event.

5 DISCUSSION

In text understanding, it is necessary to draw inferences in order to add information not presented in text explicitly. Linguistic-semantic inferences determined by semantics of a predicate can be included into the corresponding frame, e.g. as it was done in the frame of throwing in the Object role and in two frames of communicative events in the previous section. In addition, our frame structure suggests the use of a general inference scheme:

IF Act is true at time t THEN Preconditions are true at $t_1 < t$ and Consequences are true at $t_2 > t$.

In other words, if Recipient knows that Act is performed then s/he knows also which preconditions were true before performing Act and which consequences hold after performing Act.

Another and more complicated problem is representing ontological or, in general, situation-driven inferences. For instance, events proceed not always as planned by the Agent. Thus, in the sentence *John put the ball on the table but it rolled down to the floor* the ultimate location of the ball is not the table as it would be fixed in the frame of to put (in the roles Locto and Goal). Formally, this is a complex event where two motion events involving the same Object follow immediately each other and thus the ultimate location of the ball is easy to fix. When 'processing' the given sentence we understand also the connection between the events – why the rolling event occurred at all (at least supposedly). It is an example of the situation where the interaction between motion participants and motion space discussed in short in Section 3 comes into play: the rolling object must have certain shape and the ground must be relatively flat; and in the given case the top of the table was not quite horizontal, etc.

It is an example of the situation where the interaction between motion participants and motion space discussed in short in Section 3 comes into play: the rolling object as motion participant must have certain shape and the ground as motion space must be relatively flat; and in the given case the ground (top of the table) was not quite horizontal, etc.

The important fact to point out here is that the same kinds of unplanned events can occur (and can be dealt with) in communication as well. For instance, when A tells something to B with the goal that B will know the information, accept it, respond to it etc., it can happen that B does not hear A, does not understand what s/he said, or decides to not accept it or not respond to it. Analogous why-questions appear in understanding such events. Thus, a similar typology of forwarded information units and of the parameters of communication space and their interactions has to be worked out. We have already made some investigations in (Koit and Õim, 2004; Hennoste et al., 2005).

6 CONCLUSIONS AND FUTURE WORK

Texts used in human communication rely heavily on the background knowledge that the participants are supposed to have and, because of this, are not explicitly stated in text. When we are dealing with the semantic analysis of texts by the computer, this knowledge must be made explicit: it is critically relevant for constructing a coherent picture of the events, processes etc. described in a text, so that the computer would be able to fill in the 'data gaps', by taking the lacking data from its domain model, by making inferences about the event itself, about the participants, and so on.

In general, most of the motion events we are used to treat as compact and simple ones in fact

appear to be complex, especially in the context of a concrete text, and should be represented as 'compositions' of sub-events. Formally, the sub-events have the same structure as motion frames in general, but the critical requirement is that the prerequisites and consequences of sub-events, as well as the fillers of the corresponding semantic roles should fit each other in different sub-events in the way determined by the cover event.

Dialogues as social events have more complex and specific structure than physical motion events. But the general conception underlying both of them has much in common: both domains are dynamic, something is moving also in dialogues. The domain of physical motion has been studied in more detail in semantics therefore the results attained here could be used also in dealing with dialogues.

Our further work will be focused on typology of the features of the entities and their interrelations in physical motion as well as in social domain. The central aim, in studying the domain of motion, is to build a typology of entities that function as motion participants and motion spaces, and on this basis, the typology of motion events. The same type of research will be done in the domain of communication. Departing from these results, some conclusions and generalizations should be possible to make about how the process of understanding texts (and the world) is organized in humans and how these processes could be more adequately modelled on the computers.

ACKNOWLEDGEMENTS

This work is supported by the European Regional Development Fund through the Estonian Centre of Excellence in Computer Science (EXCS), the Estonian Research Council (grant ETF9124), and the Estonian Ministry of Education and Research (grant SF0180078s08). The authors thank the anonymous reviewers.

REFERENCES

Fontenelle, T. (ed.), 2003. *International Journal of Lexicography.* Special issue, 16 (3).

Hennoste, T., Gerassimenko, O., Kasterpalu, R., Koit, M., Rääbis, A., Strandson, K., Truu, T., Valdisoo, M., 2005. Miscommunication in Spoken Dialogues and Its Modelling in a Dialogue System. In *Proc. of SPECOM 2005. 10th International Conference on Speech and Computer*, 413–416. Patras, Greece.

Jackendoff, R., 2002. *Foundations of Language. Brain, Meaning, Grammar, Evolution*, Oxford University Press. New York.

Koit, M., Õim, H., 2004. Argumentation in the Agreement Negotiation Process: A Model that Involves Natural Reasoning. In *Proc. of the Workshop W12 on Computational Models of Natural Argument. 16th European Conference on Artificial Intelligence,* 53–56. Valencia, Spain,

Koit, M., Valdisoo, M., Gerassimenko, O., Hennoste, T., Kasterpalu, R., Rääbis, A., Strandson, K., 2006. Processing of Requests in Estonian Institutional Dialogues: Corpus Analysis. In *Text, Speech and Dialogue, Proceedings*, 621–628.

Lakoff, G., 1987. *Women, Fire, and Dangerous* Things, University of Chicago Press. Chicago.

Õim, H., 2012. Ontological Features of Entities in Motion Events and Their Role in the Semantic Analysis of Sentences. In *Human Language Technologies – The Baltic Perspective,* 280–285, IOS Press. Amsterdam etc.

Õim, H., Orav, H., Kahusk, N., Taremaa, P., 2010. Semantic analysis of sentences: the Estonian experience. In *Human Language Technologies – The Baltic Perspective,* 208–216, IOS Press. Amsterdam etc.

Pustejovsky, J., 1995. *The Generative Lexicon,* The MIT Press. Cambridge, Mass.

Schank, R. C., 1986. *Explanation Patterns: Understanding Mechanically and Creatively*, Hillsdale, NJ Lawrence Earlbaum Associates.

Schank, R. C., 1975. *Conceptual Information Processing,* Elsevier. New York.

Schank, R. C., Abelson, R. B., 1977. *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ Lawrence Earlbaum Associates.

Tversky, B., Zacks, J. M., Morrison, B. J., Hard, M. B., 2011. Talking about events. In J. Bohnemeyer, E. Pedersen (eds.) *Event representation in language and cognition,* 216–227. CUP. Cambridge etc.