

Semantic analysis of simple sentences: the way to go for Estonian

Neeme Kahusk

University of Tartu
Institute of Computer Science
Liivi 2 – 308, Tartu, Estonia

Nelijärve 2011

Outline

- 1 Motivation
- 2 Steps to Go
- 3 Frames and Semantic Roles

Real language understanding in a limited domain

- Real understanding of texts
- Semantic representation of what reader knows after reading the text
- Question answering
- Towards an intelligent device acting in real room

Steps to Go

- plain text
- morphological analysis and disambiguation
- syntactic analysis and disambiguation
- word sense disambiguation
- frame semantics (and disambiguation)
- inferences
- sentence generation

Steps to Go

Mary went from Tartu to Nelijärve.
Mari läks Tartust Nelijärvele.

- Where Mary was?
- Mary was in Tartu
- Where Mary is?
- Mary is at Nelijärve

Mari

Mari+0 //_H_ sg g, sg n, //

mari+0 //_S_ sg g, sg n, sg p, //

mari+0 //_S_ sg n, //

läks

mine+s //_V_ s, //

Tartust

Tartu+st //_H_ sg el, //

Nelijärvele

neli_järv+le //_S_ sg all, //

- Kus Mari oli?
- Mari oli Tartus
- Kus Mari on?
- Mari on Nelijärvel

Main problems

- inventory of semantic roles
- moving from syntactic tree to a semantic frame
- inferences
- world knowledge

Frames and Semantic Roles

- “Estonian” frames: 4 basic motion frames
 - agentive (self-)motion
 - non-agentive (self-)motion
 - agentive (causing) motion
 - non-agentive (causing) motion
- Framenet frames: many motion frames — e.g. self_motion, cause_motion, using_vechile, mass_motion ... (we are currently using 17)

AGENTIVE SELF-MOTION

HYPERONYM: MOTION

ROLE STRUCTURE

Participant Roles

AGENT (participant who controls his/her
activity,
the instigator of the event)

FRAME: ASETSEMA1 'be located'

Object: = Agent

Loc = Locfrom

Time = Timefrom

FRAME: ASETSEMA2

Object = Agent

Loc = Locto

Time = Timeto

INSTRUMENT

[the same ASETSEMA subframes attached as by AGENT
only Object = Instrument,
which means that INSTRUMENT is supposed to move
the same way as AGENT]

Loc-Roles

LOCFROM (starting place, e.g. from the garden,
from under the table, from the box)

Locfrom-in

Locfrom-at

LOC (where the motion takes place, e.g.
on the street,

in the garden, under the table)

Loc-in

Loc-at

LOCTO (the ending place, e.g. onto the street,
into the garden, into the box)

Locto-in

Locto-at

Time-roles

[The same system: TIMEFROM, TIME,
TIMETO, DURATION]

/---/

Other roles

Not important in the given context:
DIRECTON, PATH,
MANNER, about 30 in total.

Mary went from Tartu to Nelijärve

AGENTIVE SELF-MOTION

HYPERONYM: MOTION

AGENT = Mary

FRAME: ASETSEMA1 [before]

Object = Mary

Loc = Tartu

FRAME: ASETSEMA2 [after]

Object = Mary

Loc = Nelijärve

John threw ball into the box

AGENTIVE CAUSING MOTION

HYPERONYM: MOTION

AGENT = John

FRAME: ASETSEMA1 [before]

Object = John

Loc = <previous>

FRAME: ASETSEMA2 [after]

Object = John

Loc = <previous>

FRAME: ASETSEMA1 [before]

Object = ball

Loc = <previous>

FRAME: ASETSEMA2 [after]

Object = ball

Loc = box

Steps to Go

John threw ball into the box

- Where John was?
- John was here
- Where John is?
- John is here
- Where ball was?
- ball was here
- Where ball is?
- ball is in the box

Summary

- Morphological analysis and disambiguation
- Syntactic analysis and disambiguation
- Word sense disambiguation (?)
- Frames: From Framenet or our own?
- World knowledge
- Inferences