

Perkomputila tekstokompreno

ĉu pli facile en Esperanto?

Kompreni tekston

Ni povas diri
ke la sistemo “**komprenas**” la tekston,
se ĝi povas
reagi per ĝustaj respondoj al demandoj pri la teksto.
Ĉi tie la vorto “ĝustaj” signifas, ke ankaŭ homo respondus simile.

Diversaj esploroj diras pri pluraj niveloj de kompreno,
ekde plej simplaj ĝis tre intelektcaj.

Ni parolos pri kvin niveloj.

Ekzempla teksto

Je la oka horo matene, post la matenmanĝo, Maria iris al la universitato.

Je la dua horo ŝi revenis hejmen.

Post la tagmanĝo ŝi promenis en la parko kaj ĝuis freŝan aeron.

Je la kvina horo ŝi renkontiĝis kun sia amikino, kaj ili spektis novan kinon en apuda kinejo.

Poste Maria revenis hejmen kaj faris la hejmtaskon.

Nivelo 1. La teksto mem

- a) Sistemo komprenas faktojn, rekte skribitajn en la teksto.
- b) Plejparte sufiĉas morfologia kaj sintaksa analizoj. Por certaj demand-vortoj (kie, kien, kiam) necesas ankaŭ iomete da semantiko.
- c) Trovi en la teksto samajn partojn, kiel en la demando.

Nivelo 1. Ekzemplaj demandoj

D: Je kioma horo *Maria iris al la universitato?*

T: Je la oka horo matene, post la matenmanĝo, *Maria iris al la universitato.*

R: Je la oka horo.

Nivelo 1. Ekzemplaj demandoj

D: Ĉu ŝi spektis novan kinon?

T: Je la kvina horo ŝi renkontiĝis kun sia amikino, kaj *ili spektis novan kinon* en apuda kinejo.

R: Jes (konsiderante ke *ili* = ŝi kaj ŝia amikino).

Nivelo 1. Ekzemplaj demandoj

D: *Kion Maria faris post la tagmanĝo?*

T: *Post la tagmanĝo ŝi promenis en la parko kaj ĝuis freŝan aeron.*

R: Promenis en la parko kaj ĝuis freŝan aeron.

Nivelo 1. Ekzemplaj demandoj

D: *Kion Maria faris post la tagmanĝo?*

T: *Post la tagmanĝo ŝi promenis en la parko kaj ĝuis freŝan aeron.*

R: Promenis en la parko kaj ĝuis freŝan aeron.

Nivelo 1. Ekzemplaj demandoj

D: Kiam *Maria iris al la universitato?*

T: Je la oka horo matene, post la matenmanĝo, *Maria iris al la universitato.*

R: Je la oka horo.

Nivelo 2. Faktoj el la teksto

- a) Sistema komprenas faktojn, kiuj sekvas el la strukturo de la teksto.
- b) Uzatas t.n. “tekstaj logikoj”: tempa, spaca/dimensia, kaŭza, kiuj povas produkti novajn faktojn, kiuj mem ne estas en la teksto.
- c) "Tempa fadeno", "lokaj koordinatoj", matrico de statoj.

Nivelo 2. Ekzemplaj demandoj

D1: Kio okazis pli frue — iro al la universitato aŭ tagmanĝo?

D2: Ĉu Maria promenis post reveni el la universitato?

T: ... Maria iris al la universitato. ... Post la tagmanĝo ŝi promenis en la parko ...

R1: Pli frue estis iro al la universitato.

R2: Jes, ŝi promenis post reveni el la universitato.

Nivelo 3. Scioj pri la mondo

- a) Kvalitoj de objektoj; interrilatoj de objektoj; scenaroj, proceduroj, procezoj, statŝanĝoj.
- b) Uzatas faktaro (sci-bazo, simile al datum-bazo), hierarkio de objektoj kaj nocioj.
- c) Serĉado pere de semantikaj formuloj:
“Kie estis Maria vespere?” →
[“juna persono insekse” + “dum vespera tempoperiodo” + “faras ion” + “en kia ajn loko”]

Ŝi spektis novan kinon en apuda kinejo. Hejme ŝi faris la hejmtaskon.

Nivelo 3. Ekzemplaj demandoj

D: Ĉu ŝi studis je tagmezo?

A: *Necesas scii ke universitato estas loko kie oni studas, kaj ke tagmezo estas sama tempopunkto kiel la dekdua horo.*

R: Jes. (Ŝi studis en la universitato).

Nivelo 3. Ekzemplaj demandoj

D: Je kioma horo Maria revenis hejmen vespere?

A: *Konsiderante ke kutime kinospektado daŭras proksimume horon kaj duonon aŭ du horojn.*

R: Ĉirkaŭ la sepa horo.

Nivelo 4. Dialogo kaj rezonado

- a) Konsidero de diskurso kaj dialogo.
- b) Klarigoj pri “kial vi decidis respondi tiel?”.
- c) Eĉ tiaj ekstremaĵoj kiel edukiteco de la tekst-aŭtoro kaj de la demandanto.

Nivelo 5. Lernokapablo

- a) Konsidero de antaŭe legitaj tekstoj kaj dialogoj.
- b) Pliĝustigo per kromdandoj.
- c) Pliriĉigo de la faktaro.
- d) Abstraktigo, ĝeneraligo kaj detaligo.
- e) Trovo de samaj informoj, decidoj pri malsamaj informoj (diversaj opinioj).
- f) Fidebleco de tekstofontoj.

Tekstoproduktado

- Por respondi la sistemo devas kapabli mem produkti tekstojn.
- Plej komence, eblas disponigi al la sistemo respond-variantojn, kaj ĝi elektu taŭgan.
- Sur nivelo 1: reprodukto de tekstopartoj.
- Sur nivelo 2: iom pli komplikita reprodukto.
- Sur nivelo 3: traduko el semantika lingvo al la homa lingvo (Esperanto niakaze).

Stumbligiloj kaj defioj

- eraroj en la tekstoj
- kiel prezenti sciojn en komputilo (strukturoj)
- kiel instrui la programon pri novaj scioj
- kiel priskribi la lingvon
- kiel priskribi la mondon kaj ties trajtojn
- multsignifaj vortoj
- por kiu celo ni aplikas la sistemon
- ĉu (iu) teksto estas en Esperanto?
- formala teorio de dialogo inter homo kaj maŝino

Kompreno per komputilo

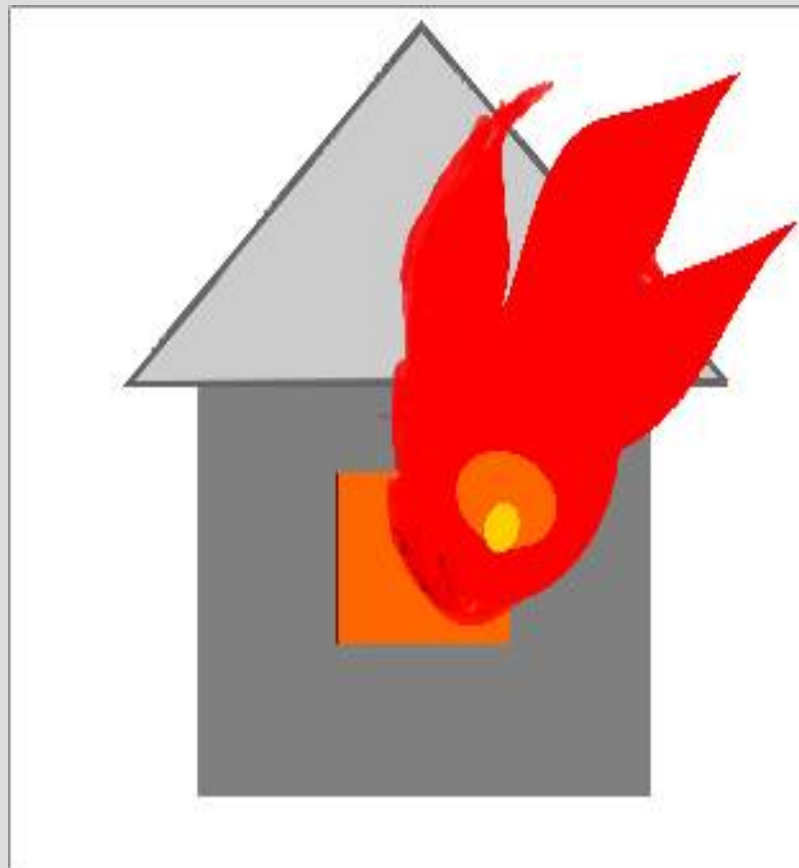
"Komputilo nenion povas kompreni,
ĝi estas nur metalaĵo.
Komputilo povas nur komputi kaj reagi."

V. Tuzov,
aŭtoro de semantika vortaro
de la rusa lingvo

Domo brulas!

"Nu, kaj kion mi faru?", — pensas la komputilo.

a) bildigi tion?



Domo brulas

voki fajrobrigadon
aŭ / kaj
funkciigi estingilojn

(reagi al evento)

Domo brulas

trovi similan situacion:

en certa tekstaro,
en dokumentaro,
en bildoj

Domo brulas

foti ĝin

kaj

publikigi ĉe Facebook

Domo brulas

reagi emocie

(tasko por roboto)

Domo brulas

diveni la kaŭzon

(nu, post multaj jaroj verŝajne)

Domo brulas

fari nenion

(eĉ tiel, jes)

Dependas de la taskoj!

Sed la komuna afero en ĉiuj ĉi situacioj estas ke ni povas prezenti la frazon al komputilo kiel funkcion *brulas1(domo)*.

Por alia frazo, ekz. “brulas vangoj” =>
brulas2(j(vango)).

Tia transformigo estas fakte formaligo de lingvo.

Formaligo de lingvo

1. **semantika lingvo** por priskribi sencojn (pli-malpli sendependa de lingvo);
2. **semantika vortaro** por priskribi vortojn de konkreta lingvo (sencojn kaj valentojn);
3. **semantika analizilo** por transformi tekston al funkciaj strukturoj kaj reen.

Surbaze de teorio kaj praktiko de semantika vortaro de la rusa lingvo fare de profesoro Vitalij Tuzov.

Ekzemplo

"Mi aĉetis libron de Petro" →

@frazo aĉetis(@kiu mi, @kion libron, @de_fonto de(@kiu Petro))

aŭ

@frazo aĉetis(@kiu mi, @kion libron(@kies de(@kiu Petro)))

Vortoj kiel funkcioj

@verbo_transitiva_is **aĉetis**(*@kiu*, *@kion*, *@kie*, *@de_fonto*, *@je_prezo*, ...)
@kio **libro**(*@de_aŭtoro*, *@pri_temo*, *@ecoj*, *@kies*, ...)

@frazo **kaj**(*@frazo*), kiel en frazo "kaj ili foriris"

@verboj **kaj**(*@verbo1*, *@verbo2*)

@celo **por**(*@verbo_i*)

@celo_persono **por**(*@kiu*)

@celo_afero **por**(*@kio*)

@kies **de**(*@kiu|@kio*)

@de_fonto **de**(*@kiu|@kio*)

@de_punkto **de**(*@loko*)

@de_aŭtoro **de**(*@kiu*)

@de_aganto **de**(*@kiu*), eble ĉi tie taŭgus ankaŭ *@kio*

Vortoj kiel funkcioj

Kaj jen kiel tio funkcias (simpligite): jen ni havas sintagmon “de Petro”, kiu havas variantojn:

@de_aganto **de**(*@kiu Petro*)

@kies **de**(*@kiu Petro*)

@de_aŭtoro **de**(*@kiu Petro*)

Se ni provus ligi ĝin al la vorto “farita”:

@participo_pasiva_it **farita**(*@celo, @de_aganto, @periodo, @momento, @loko, ...*)

la sistemo elektus la unuan varianton, ĉar ĝia tipo *@de_aganto* estas en la argument-listo de la vorto “farita”, kaj por la vorto “libro” ni elektus la duan kaj la trian variantojn (do devus pliĝustigi, ekz. per demando aŭ kunteksto).

Aplikebloj

- resumoj de tekstoj (temo, herooj, agoj, tempo(j) de la eventoj ktp)
- inform-retrovado (trovi simil-temajn tekstojn aŭ tekstojn pri samaj objektoj, personoj)
- respond-sistemoj (surbaze de certa tekstaro, anticipe indeksita semantike)
- memlernado de fakt-bazoj (per analizi donitan aron de tekstoj aŭ per dialogo kun spertuloj)

Dankon!

Konstantin Hlyzov,
Irkutsk, Rusio
2010