

SZEMANTIKA, 1. ELŐADÁS

Kornai András

BME 2020/9/9

MI A SZEMANTIKA?

- A jelentés tudománya
- Jó, de mi a jelentés?
- Amit megértünk (vagy nem)
- Jó, de mi az hogy érteni?
- Erre a kérdésre már lehet válaszolni!
- McCarthy (1976) példája

NEW YORK TIMES 1973 SZEPT 29

Két rabló egy teherlift aknájába lőkött egy 61 éves bútorkereskedőt tegnap a Brooklyn belvárosában található boltjában, miközben a harmadik rabló megpróbálta a lifttel agyonnyomni. A rablók nem voltak elégedettek azzal, hogy csak 1200 dollárt tudtak belőle kikényszeríteni.

A kereskedőt, John J. Hugot lelökték a földszintről a pincébe, de az akna alján található fékezőrugók megakadályozták, hogy a lift agyonnyomja. Körülbelül 30 centiméterrel felette állt meg a lift, míg ő a gödör aljához lapult. Hug körülbelül fél órán át volt beszorulva az aknába, míg végül a kiáltásaira felfigyelt egy szállítómunkás. A Livingston Street 340 alatti bolt a Seaman's Quality Furniture üzletlánchoz tartozik.

Hugot rohamrendőrök mentették ki, és a Long Island College Kórházba szállították. Sokkos állapotban volt, de miután ellátták a horzsolásokat a bal karján és a gerinctájon található sérüléseit, elengedték, és hazament. A 69th Lane 62-01-ben lakik, Maspethben (Queens).

Hét éve dolgozott a boltban, a Nevins Street sarkán. Azóta negyedszer fordult elő, hogy rablók támadták meg a boltban. Utoljára egy éve, akkor a jobb karját vágta meg egy késsel felfegyverzett betörő.

McCARTHY (1976) SZERINT

A történetben leírtak alapján bármilyen értelmes személynek vagy programnak választ kellene tudnia adni a következő kérdésekre:

- 1 Ki volt a boltban, amikor az események megkezdődtek?
- 2 Ki volt a boltban akkor, amikor megpróbálták megölni Hugot?
- 3 Kinél volt a pénz a végén?
- 4 Él-e ma Hug?
- 5 Hogyan sérült meg Hug?
- 6 Hol lakik Hug?
- 7 Brooklynban lakik-e Hug?
- 8 Mi a rablók neve és címe?
- 9 Eszméletén volt-e Hug, miután a rablók távoztak?
- 10 Mi történt volna, ha Hug nem lapult volna az akna aljához? Mi történt volna, ha nem lettek volna fékezőrugók?
- 11 Akarta-e Hug, hogy agyonnyomja a lift?
- 12 Bemutatkoztak-e a rablók Hugnak?

MI A SZEMANTIKA CÉLJA?

- Megérteni, hogy valami mit jelent
- Jó, de mi az a valami?
- Lehet egy jelzőtábla, számítógépes program, beszélt/írott szöveg, ...
- A megértés kritériumai:
 - ▶ Tudni a dologgal kapcsolatos kérdésekre értelmesen válaszolni
 - ▶ Tömören összefoglalni
 - ▶ Meggyőzően folytatni a szöveget (GPT-3)
 - ▶ Átfogalmazni (parafrazálni)
 - ▶ Más nyelvre lefordítani
 - ▶ Matematikai modellt adni “megértettem a gravitációt”
- Az emberek képesek minderre “vannak szemantikai képességeik”
- A szemantika célja az, hogy a gépeket is képessé tegyük

A LEGFONTOSABB TECHNIKAI FOGALOM

A JEL (SAUSSURE 1916)

A jel egy forma és egy jelentés egysége. A formát hívjuk *jelölőnek* (signifier) a jelentést *jelölt-nek* (signified, meaning) is. 'Egység' helyett ma inkább rendezett párról beszélünk.

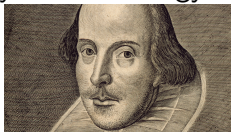
- Ebben a kurzusban nem sokszor lesz szó véges jelhalmazokról (pl. közlekedési jelzőtáblák), mert itt a szemantikai képességek memorizálással pótolhatók
- Sokszor lesz szó szavakról (végtelen sok szó van!), mondatokról, és nagyobb szövegekről
- A természetes nyelvre koncentrálnak (van olyan, hogy "a programnyelvek szemantikája" de itt nem sok szó lesz róla)
- Vannak komoly tesztek, különösen a GLUE és a SUPERGLUE

HOGY LEHET A JELENTÉST FORMALIZÁLNI?

A JELENTÉS

Ez nyilván valami az emberek fejében, de abba nem látunk bele.
Tehát: formális modelleket készítünk

- Annyit azért tudunk, hogy az agyban neuronok vannak. Innen származnak, **de nem direkt anatómiai úton!** az *ideghálós modellek*
- Azt is tudjuk, hogy a szavak nagy része dolgokra referál, pl.



Shakespeare erre: . Innen erednek a *logikai modellek*

- Azt is tudjuk, hogy a mondatoknak van egyfajta belső struktúrája, innen erednek a *tudásreprezentációs modellek*
- A kurzus alapján ezekről (és kapcsolataikról) szól majd

IRODALOM

- Kötelező a mai órához:
<https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion>
(GPT-3)
<https://bit.ly/2EI9ONI> (How many words are there?)
<https://w4ngatang.github.io/static/papers/superglue.pdf>
(Superglue)
- Ajánlott a továbbiakban:
https://www.tyopotex.hu/book/9642/kornai_andras_szemantika
(Kornai: Szemantika)
<https://www.springer.com/gp/book/9783319656441>
(Kornai: Semantics)
<https://kornai.com/Drafts/advsem.pdf> (Advanced Semantics)
- A kurzus honlapja: <https://kornai.com/2020/Semantics>

LEGKÖZELEBB: JELENTÉS ÉS INFORMÁCIÓ

- Mi az az információ, mi köze a jelentéshez?
- Hol van az információ? (Protip: nem a fejben)
- Szorgalmi 1: gyűjts össze sok szöveget (legalább 1m szót)
- Egy nyelvből, de az az egy ne legyen se angol se magyar
- Számláld meg melyik szó hányszor fordul elő, legyen az i -edik leggyakoribb szó abszolút gyakorisága F_i
- Ábrázold F -et log-log skálán