



Itthon Világ Gazdaság IT | Tudomány Panoráma Vélemény Kultúra Cégautó Vállalkozás Egészség Karrier

HVG > 2007/22. szám

 SZÁMÍTÓGÉPES FORDÍTÓPROGRAMOK

A szavak embere

2007. május 30. 00:00 | Utolsó módosítás: 2007. május 30. 17:37

Prószéky Gábort, a nyelvészeti programok specialistájának számító MorphoLogic ügyvezetőjét 1978 óta foglalkoztatja a számítógépes fordítóprogramok világa. Akkor egy nyarat adott magának egy ilyen rendszer létrehozásához. Ma viszont úgy látja, a gép sohasem lesz képes hibátlanul visszaadni az idegen nyelven írt szövegeket, de mégis szükség van az egyre jobb megértést támogató programokra.

Háló Komolyan gondolta társaival 29 éve, hogy egyetlen nyár leforgása alatt képesek lesznek gépi fordítóprogramot írni?

Prószéky Gábor: Igen, és őszintén meglepődöttünk, hogy nem sikerült. Abban a hitben éltünk, hogy mindent tudunk programozni, miért ne tudnánk hát írni egy emberi nyelveket fordító programot. Kiderült, hogy ez nem is olyan egyszerű. Ma már tudom, nem lesz olyan nyárvége, amikor egy tökéletes fordítói programot el lehetne készíteni.

Háló Mi a gond a gépi fordítással?

P. G.: A teljes kontextust képtelenség feldolgozni, lehetetlen a jelentésárnyalatokat "gépileg" megkülönböztetni, erre az emberek által írott szótárak sem vállalkoznak. Csak a sokadik jelentésnél adják meg, hogy az jogi vagy mondjuk pejoratív jelentésű, ám a szótárak éppen a leggyakrabban használtknál hallgatnak. Gyakorlatilag nincs olyan, hogy "szó": az mindig csak a környezetéből következik. Olyan sincs, hogy szinonima, mert például a kutyát "általában" használhatom eb értelemben, de csak ebadóról, nem pedig kutyaadóról beszélhetek. Hiába tudnak a fordítóprogramok több százezer szabályt, annál az ember többet vagy mást használ. Különösen az okoz gondot, hogy nem tudjuk, mire használják majd a programot. Autóműszaki ügyeknél például másra van szükség, mint a sportnyelvben. Ha sikerülne specializálni a programokat, akkor sokkal kevesebb lenne a szerencsétlen szóválasztás.

Háló Az interneten megjelentek www.webforditas.hu alkalmazásukkal, amely ingyenesen kínálja angol nyelvű weblapok magyar fordítását. A maga suta eredményével is megkapó, hogy a program egyetlen pillanat alatt "lefordítja", mondjuk, a BBC portáljának címlapját.

P. G.: Ezt a honlapot cégünk 15 éves évfordulójára indítottuk el tavaly. Igazából nem is fordítást adunk, hanem az idegen nyelvi szöveg megértését támogatjuk. Félő, hogy az emberek félreértik a fordítás fogalmát. A fordítás ugyanis hagyományosan azt jelenti, hogy van egy forrásnyelvi szövegem, amiből előállítják a célnyelvi szöveget. Ám a legtöbb esetben nincs is szükség erre a "tökéletes" transzformációra. A Tihanyi László kollégám által vezetett kutatás - a MetaMorpho fordítórendszer - eredménye többféle "ruhában" is elérhető: a MoBiCAT - a jól ismert MoBiMouse-hoz hasonlóan - a szövegre való rámutatással működik, de ez már teljes mondatokat fordít. Abban is különbözik az elődjétől, hogy nem lokális vagy hálózati eszköz, hanem egy - az előfizetőink kéréseit figyelembe vevő - állandóan finomodó internetes szolgáltatás. Új eszközünk, a Word szövegszerkesztőbe integrált MorphoWord fordítóprogram - és a korábban említett [webforditas.hu](http://www.webforditas.hu) is - már képes arra, hogy az angol szöveget teljes egészében magyarra ültesse át, s éppen az elkövetkező napokban jelenünk meg ezek magyar-angol változatával is. Ez utóbbinak az adatbázis-építésébe bekapcsolódott az MTA Nyelvtudományi Intézete és a Szegedi Tudományegyetem is. Az angol és bármilyen más nyelv közötti fordítás így könnyebben megvalósítható, hiszen a munka egy jelentősnek mondható részét már elvégeztük.

Háló Úgy tudni, valami egész mást csinálnak, mint a többi gépi fordító cég. Milyen elvet követnek?

P. G.: Egy mondatban minden szó csak azzal van ott, amit a többi megenged neki, a környezet alapvetően meghatározza a szó jelentését. Ez a szemlélet elvezethet a mondat tartalmának - hozzátétőleges - megértéséhez. 2000-ben jutottunk el oda, hogy komolyan vehessük az örök szerelem, a gépi fordítóprogram elkészítését, ugyanis akkorra álltak össze az addigi fejlesztéseink, a nyelvhelyességi rendszer, a szótóvesítő, illetve a szótári és mondatelemzési kutatásaink és fejlesztéseink.

Alapvetően a gépi fordításnak két módszere létezik. Az újabb módszer - statisztikai eljárásokkal - elképesztő anyagmennyiséget tud kezelni. Végtelen mennyiségű szövegmintha esetében ez lenne a tökéletes megoldás, de ma még kevés a feldolgozható szövegyanyag ahhoz, hogy jó minőségű fordítás készülhessen. A régi módszerrel meg valahogy úgy vagyunk, mint a varánuszgyíkkal, azaz ki kellett volna pusztulnia, de a Galápagos-szigeteken mégis életben maradt. Egy-két olyan gépi fordítói rendszer él a hatvanas évek óta, amely többször "átvedlett", ám mivel hihetetlen mennyiségű munka van bennük, a mindenkori tulajdonosnak észébe se jut kidobni őket. Ugyanakkor fejlesztésükre már nem érdemes pénzt és időt áldozni.

A mi módszerünk valahol a kettő között van, mintákat is használunk, noha azok nem statisztikaiak. Megpróbálunk minél nagyobb egységeket kezelni: a legtöbb fordítórendszer a szó szintjéről indul, mi meg nem. Valahogy úgy állunk ehhez, mint a nyelvtanárok: tessék kifejezéseket, frázisokat tanulni, ne pedig szavakat - s ha ez nem megy, csak akkor lépünk tovább alacsonyabb szintre, például a szavakéra. A legtöbb fordítás úgy dolgozik, hogy előbb elemez, ezt követően átalakítja, majd "kigenerálja" a szöveget. Nálunk már az elemzés ideje alatt készen állnak a megoldások. A kétszázézer, általunk alkalmazott jelenlegi szabály még igen kevés. Egy példával illusztrálnám ezt: a német szinkrontolmács úgy dolgozik, hogy "megelőlegezi" a mondat állítmányát, pedig a fránya német szórend miatt az igen sokszor csak a mondat végén hangzik el. A MetaMorpho rendszer is ezt teszi, mert részszerkezetből rakja össze a szövegajánlatát, és amikor az elemző eléri a mondat végét, kész is a fordítás.

Háló Ön is csak valószínűségekről beszél. De azért egy Kosztolányi-versbe biztosan beletörne a program bicskája...

P. G.: Ezt a kérdést nem szeretem. Egy gépi fordítói program sosem lesz képes szépirodalom fordítására, mint ahogy az emberek nagy része sem. Ahogy Nádasdy Ádám csipkelődve szokta mondani: a műfordítás "tönkretezi" a pontos fordítást, hiszen a műfordító számtalan esetben a pontos fordítás fölé rendel más szempontokat, például hogy ki kell jönni a szótagszámnak, vagy egyszerűen csak egy adott kor ismerete másfajta megfogalmazást kíván. Egyébként az újságok politikai rovatát sem tudja a gép sokszor jól lefordítani, de egy autó vagy egy számítógép műszaki leírásának átadására előbb-utóbb képes lehet.

Háló Mire lehet akkor most használni a fordítóprogramokat?

P. G.: Mondok egy példát: a magyar gazda ma nem kap meg a tevékenységére vonatkozóan minden EU-s információt magyar nyelven. Ha viszont átnézi az unió megfelelő honlapjait a mi fordító-, pontosabban megértést segítő programunkkal, már könnyebben tájékozódhat arról, hogy például kölest vagy árpát vessen-e. Időt és pénzt takarít meg, mert maga tudja kiválasztani, hogy melyik angol nyelvű oldalra van szüksége, amit azután pontosan lefordíthat egy hozzáértővel.

Háló Milyen az angol-magyar fordítójuk fogadtatása?

P. G.: Az elemzések azt mutatják, hogy miközben a magyarok szenvedélyesen elutasítják a gépi fordítás zagyvának tűnő "eredményét", mégis nagyon jól megértik az így fordított szövegeket. Programunk nyelvi segédeszköz, éppúgy, mint a helyesírás-ellenőrző. Én sem szeretek mankóval járni, de megértem, hogy másnak nagy szüksége lehet rá, mert egyébként nem tudna haladni.

Háló Magyarországon az emberek nyolcvan százaléka egyáltalán nem beszél idegen nyelveken. Nem teszi védtelenné a nyelvet nem beszélőt egy ilyen program? Még azt hiheti, hogy minden rendben, mert a gép úgymint ügyesen lefordít majd mindent.

P. G.: Érdekes jelenség, hogy amikor a kontroll kikerül a felhasználó kezéből, nagyobb lesz az elégedettség és egyben a veszély is. Azt látja ugyanis az angolul nem tudó magyar, hogy zagyván ültették át az angol szöveget magyarra, de könnyedén képes arra, hogy érthetővé alakítsa át magának a zavaros mondatokat. Ám az általa nem beszélt angolra fordított szövegre már nincs befolyása. Ezért létrehoztunk egy úgynevezett kontrollált nyelvi eszközt, amelynek az a célja, hogy bizonyos bonyolult nyelvi szerkezetek ne jelenhessenek meg a fordítandó szövegben, mert azokat nem képes a program angolra átültetni. Figyelmeztetjük a felhasználót, hogy például nem írhat hosszan és bonyolultan: ha a szabályokat betartja, akkor nagy valószínűséggel értelmes angol fordítás készül a szövegéből.

ZÁDORI ZSOLT