

HOZZÁSZÓLÁS

"A FORDÍTÓGÉP TÜNDÖKLÉSE ÉS BUKÁSA"

C. CIKKHEZ

Makay Árpád

József Attija Tudományegyetem
Kibernetikai Laboratórium, Szeged

Erdemes összehasonlítani a "hagyományos fordítás" és a "gépi fordítás" jellemzőit.

A "hagyományos" fordító ember, aki egy idegen nyelv elsajátítását legalább egy másik nyelv anyanyelvi szintű ismeretében kezdi. Az "anyanyelvi szint" nemcsak nyelvtani kategória, hanem a valóság olyan mértékű ismeretét is feltételezi, amely biztosítja a nyelvi egységek /szavak, mondatok stb./ valóságelemekre való helyes alkalmazását is, azaz a jelentést. Az anyanyelvi egységek jelentéseinek, így magának a nyelvnek az ismerete összes érzékszerveink és gondolkodásunk produktuma. Ezzel szemben egy idegen nyelv elsajátítása sokkal kisebb apparátus felhasználásával is lehetséges /nem is szólva arról, hogy esetleg csak az idegen nyelven leírt szöveg megértését tűzzük ki célul/. Ez annak a következménye, hogy az idegen nyelvi egységek jelentéseit az /ismert/ anyanyelvi jelek jelentéseire /és nem a valóságra/ vezetjük vissza.

Ezek után az idegen nyelvre vagy nyelvről való fordítás anyanyelvről, illetve anyanyelvre a megtanult jelentés-azonosságok alapján történik, és így természetes módon eleget tesz a fordítás azon alaptörvényének, hogy egy A nyelvi mondat megfelelőjének mindazokat a B nyelvi mondatokat tekintjük, amelyek jelentése azonos az eredeti mondattal.

Tegyük azonban hozzá egyrészt azt, hogy vannak olyan valóságelemek, amelyekről először az idegen nyelv tanulása során hallunk és közvetlenül az idegen nyelvi "képét" rögzítjük /ezt fejleszti tovább az audio-vizuális oktatás, amely az anyanyelv tanulási folyamatát igyekszik megvalósítani/; másrészt a gondolkodás ugyanugy "aktivizálódik", mint az anyanyelvi tanulás során, így soha nem hallott kifejezéseket /szótár nélkül!/ megértünk, sőt lefordítunk.

Most nézzük a gépi fordítást. A számítógép összehasonlíthatatlanul rosszabb kiindulópontból indul. Van ugyan "anyanyelve" /a gépi kód vagy valamely magasabb szintű algoritmikus nyelv/, de ez jóval szegényebb egy természetes nyelvnél. A /természetes nyelvi "közlemény"-nek megfelelő program szerepe egy eljárás /algoritmus/ közlése a számítógéppel, amely az elemi lépéseknek megfelelő utasítások egy sorozata. Az utasítások és így a program jelentése egyértelműen definiálható azzal az állapottal, illetve állapotváltozással, amelyet a számítógépben, mint rendszerben okoz.

A számítógép számára tehát bármely természetes nyelv "idegen nyelv". Ha meg akarjuk tanítani valamely természetes nyelvre, ugyan-
 úgy mint az ember esetében, le kell írunk a nyelvi egységek jelenté-
 seit a gépi kódú utasítások jelentéseinek segítségével, ami összeha-
 sonlíthatatlanul bonyolultabb, mint egyik természetes nyelv leírása
 egy másik természetes nyelv segítségével. Ehhez elsősorban a nyelvtu-
 domány szólhat hozzá tisztázva a jelentés fogalmát, másodsorban a ma-
 tematika nyelvi modell felállításával és formalizálásával, harmadik-
 nak a számítástudomány a modell működéséhez szükséges algoritmusok
 felállításával. /A programozás ezekhez képest elhanyagolható, az adat-
 előkészítés pedig az üzemeltetés költsége./

A gépi fordítás szakemberei nem tűzték célul ezt az egyelőre még
 túlságosan bonyolultnak látszó feladat megoldását, mivel egy A nyelv-
 ről B nyelvre való fordítás során "csak" a jelentés-azonosság megál-
 lapítása szükséges. Más szóval közömbös az, hogy mi a jelentés. Első-
 sorban azért választanak műszaki-tudományos szövegeket gépi fordítá-
 si kísérletek anyagául, mert egy adott szakterületen az esetek több-
 ségében a szavaknak egyértelmű azonos jelentésű párja van. A lényeges
 probléma tehát a szintaktikai analízis és szintézis, ami már a közel-
 jövőben sem látszik reménytelenül megoldhatatlannak. /Megjegyzendő,
 hogy a strukturális nyelvészet feladata a szintaxis leírása, a mate-
 matikai nyelvészet módszereket, modelleket, algoritmusokat szolgál-
 tat és segít a szintaxis formalizálásában; ugyanez a feladata a ma-
 tematikának a többi tudományág területén, tehát semmiképpen sem ki-
 ván a matematika a "világ lényegévé" válni./

A számítógépek korábbi fejlettségi fokán a gépi fordítás szük-
 ségképpen "batch-processing" üzemmódban történt, azaz az előkészített
 forrásnyelvi szöveget, mint input-ot a számítógép emberi beavatkozás
 nélkül lefordította és kiadta a célnyelvi szöveget, amelyet esetleg
 utószekesztettek számítógép nélkül. Az utóbbi évek ugrásszerű fejlő-
 dése lehetővé tette az ember és gép interaktív kapcsolatát: a fordí-
 tás során az embernek lehetősége van az eljárás módosítására, az ered-
 mények menet közbeni korrigálására, de a számítógép is "kérdezhet".
 Az interaktív üzemmód elhárítja az olyan /batch-processingben megold-
 hatatlan/ hiányosságokat, amelyek pl. egy szónak a gépi szótárból va-
 ló hiányzásából következnek. A gép "tanul": egy szituáció, amely ko-
 rábban emberi beavatkozást igényelt, később már "ismert" a gép előtt.
 Végül gazdaságos is: esetleg több száz felhasználó dolgozik egyidejű-
 leg ugyanazzal a fordító-programmal.

Nagyon szerények vagyunk akkor, ha ezt a következtetést szűrjük
 le: a számítógép eszköz, ha nem is pótolja a "hagyományos" fordítást,
 használatáról nem mondhatunk le nemcsak a fordítás, hanem a nyelvtu-
 domány, informatika stb. területein sem. A nyelv szerepe a társadalom-
 ban kulcsfontosságú: pont a nyelvet ne közelítsük meg a matematika és
 a számítástechnika módszereivel és eszközeivel?

