

V. P. B E R K O V - B. A. J E R S O V :

Kísérletek fordítógép készítésére ①

[Több olvasónk kérésére közöljük az alábbi cikket.

- A szerk.]

P 8496

1954. január 7-én New-Yorkban, az IBM cég üzletében bemutatták az IBM-701 típusu elektronikus számológépet, amelynek segítségével orosz nyelvről angolra fordíthatók szövegek. A bemutatón bebizonyosodott, hogy a gép képes helyes fordításra. Egy-egy mondat lefordításához 6-8 másodpercre van szüksége. A nyilvános próbán körülbelül 60 mondatot fordítottak le vele a politika, a kémia, a matematika, a jogtudomány, a hadügy stb. tárgyköréből.

Persze egyáltalán nem volna helyes azt hinnünk, hogy a valamely nyelvről egy másik nyelvre gépi erővel való fordítás problémája teljesen megoldódott. Az ilyen fordításnak egyelőre csak az e l v i l e h e t ő s é g e tekinthető bebizonyítotttnak. Az első kísérlet során a gép csak egyszerű bővített mondatokat fordított, amelyeket összesen 250, előre kiválogatott szóból szerkesztettek meg, s a kiválasztott szavaknak legfeljebb két orosz megfelelőjük volt megadva. Jelenleg - miként a munkálatok egyik vezetője, Leon Dostert (George Town-i egyetem) kifejezte - "még nem tartunk ott, hogy a gép egyik végén betesszük az orosz könyvet, s a másik végén kipottyan az angol".

Hogyan működik a fordítógép? Működési alapelvét olyan "tolmács"-nak a munkájához hasonlíthatjuk, akinek valamely nyelvről le kell fordítania valamit egy másik nyelvre, bár egyik nyelvet sem ismeri.

①

A "Voproszű jazükoznanija" szerkesztőségi megjegyzése: Folyóiratunk egyik közeli számában külön cikket közlünk majd a Szovjetunióban folyó fordítógép-kísérletekről is.

Lássuk a legegyszerűbb esetet. "Tolmács"-unknak a general-major, 'vezérőrnagy', orosz kifejezést kell átültetnie angolra. Legelsősorban - ez nyilvánvaló - szótárt kell adnunk neki, amelyben megtalálhatja e kifejezés minden tagjának az angol megfelelőjét. Ez azonban nem elegendő, mert amikor a "tolmács"- lépésről-lépésre haladva - az orosz general szót az angol general szóval, a majort pedig a majorral fordítja, a general-major kifejezést kapja, holott angolul az major general. Világos: a "tolmáccsal", a szavakon kívül azt a szabályt is közölnünk kell, hogy mikor kell megfordítani a szórendet. De ebből a célból a szótárban minden orosz szó mellett ott kell állnia - az angol ekvivalensen kívül - egy megfelelő egyezményes jelnek, "kulcsszám"-nak is, amely megmutatja, hogy melyik szabályhoz kell folyamodni. Az említett példa lefordításához a szótár ekkor így fog festeni:

orosz szó	angol megfelelője	kulcsszám ^②
<u>general</u>	<u>general</u>	12
<u>major</u>	<u>major</u>	.

A szabály így hangzik: "Ha az adott szó kulcsszáma 12, akkor megállapítandó a következő szó kulcsszáma. Ha ez +, akkor a szórend megfordítandó. Az ellenkező esetben a szórend változatlan marad." E szabálynak és a kulcsszámokkal ellátott szavaknak az alaki felhasználásával megkaphatjuk a helyes angol fordítást: major general.

Lássunk egy olyan példát, amikor valamelyik orosz szónak két angol megfelelője van, s a "tolmácsnak" ki kell választania a helyes megfelelést. Ez szintén a kulcsszámokkal és szabályokkal felszerelt szótár segítségével hajtható végre. Két-két főnév és előljáró szókapcsolatát kell lefordítani: otnosenie k... 'viszony vkihez' és ljubov' k... 'szeretet vki iránt'. Elkészíti a szótárt:

[lásd a következő oldalon]

② A kulcsszámok teljesen önkényesen vannak kiválasztva, s a dolog lényegén nem változtatnak.

orosz szó	angol megfelelője		kulcsszám
	I	II	
<u>otnosenie</u>	<u>relation</u>	-	111
<u>ljubov'</u>	<u>love</u>	-	48
<u>k</u>	<u>to</u>	<u>for</u>	99

és megfogalmazza a szabályt: "Ha az adott szó kulcsszáma 99, akkor megállapítandó az előző szó kulcsszáma. Ha ez a kulcsszám 111, akkor a 99-es kulcsszámú szó első megfelelője vendő, az ellenkező esetben a második. A szórend mindkét esetben változatlan marad." E szabálynak és a fenti szótárnak alaki felhasználásával megkapjuk a fordítást: relation to... és love for ...

Az teljesen nyilvánvaló, hogy összetett szókapcsolatok vagy mondatok lefordításához sokkal több szabályra és bonyolult kulcsszám-rendszerre van szükség.

Könnyen beláthatjuk, hogy az imént említettekhez hasonló kulcsszámú szótárnak és szabályoknak a segítségével a fordítás szerfölött nehézkes és hosszadalmas munka, amely egyúttal sok-sok apró-cseprő művelettel jár: a szavaknak a szótárban való keresgélésével, majd a megfelelő szabályok alkalmazásával. E teljesen gépies munkát eredményesen elvégezheti helyettünk az elektronikus számológép.

A különféle irodai számológépektől eltérően az elektronikus számológép képes teljesen önműködően elvégezni önmaga által már korábban meghatározott művelet-sorozatot is. A fordításra felhasznált IBM-701 típusu elektronikus számológépnek bonyolult szerkezete van. Elég, ha annyit mondunk, hogy 11 géptömbből áll, és teljes súlya körülbelül 10 tonna. Másodpercenként több mint 16000 összeadást vagy kivonást, s ugyancsak másodpercenként több mint 2000 szorzást vagy osztást képes elvégezni. E gépet főként a sok munkát fölemésztő atomfizikai számítások, a rakétalövedékekkel kapcsolatos számítások stb. elvégzésére használják.

Az elektronikus számológépnek van egy ugynevezett "emlékező" szerkezete, amely nagyjából a magnetofonhoz hasonlítható. Az emlékező szerkezetben megőrződő adatokat a gép körülbelül úgy keresi vissza, mint a telefonszámokat az önműködő telefonközpont. Az emlékező szerkezetben az adatokon (a jelen esetben: a kulcsszámokkal ellátott szótáron) kívül egy "programvezérlő" is található, vagyis azoknak a műveleteknek a sorrendjét megszabó szerkezet, amelyeket a gépnek el kell végeznie. A programvezérlőt ugynevezett utasítások formájában viszik át az emlékező szerkezetre. A szóban forgó esetben a

programvezérlő hat szabály (lásd alább) alapján volt összeállítva, s körülbelül 2400 utasítást foglalt magában. Megjegyezzük, hogy a legegyszerűbb mondatok lefordításához is olyan vezérlőre van szükség, amely hozzávetőleg két és fél-szer annyi utasítást foglal magában, mint az, amely az irányított lövedék pályájának számításaihoz kell.

Bemutatunk néhány mondatot, amelyet az IBM-701 lefordított:

Sorszám	A lefordítandó orosz mondat	A gép adta fordítás
1	<u>Kacsesztvo uglja opredeljaetszja kalorijnoszt'ju</u> 'A szén minőségét a kalóriatartalom szabja meg'	<u>The quality of coal is determined by calory content</u>
2	<u>Krahmal vürabatüvetszja mehanicseszkim putem iz kartofelja</u> 'A keményítőt mesterséges uton állítják elő burgonyából'	<u>Starch is produced by mechanical methods from potatoes</u>
3	<u>Obrabotka povüsaet kacsesztvo nefti</u> 'A feldolgozás javítja a nyersolaj minőségét'	<u>Processing improves the quality of crude oil</u>
4	<u>Dinamit prigotovljaetszja himicseszkim proceszszom iz nitroglicerina sz primeneniem interesznüh szoedinenij</u> 'A dinamit kémiai folyamattal nitroglicerinből készül inert vegyületek felhasználásával'	<u>Dynamite is prepared by chemical process from nitroglycerine with admixture of inert compounds</u>
5	<u>Mezsdunarodnoe ponimanie javljaetszja vazsnüm faktorom v resenii politicseszkih voproszov</u> 'A nemzetközi megértés fontos tényező a politikai kérdések megoldásában'	<u>International understanding constitutes an import factor in decision of political questions</u>

Sor-szám	A lefordítandó orosz mondat	A gép adta fordítás
6	<p><u>Velicsina ugla opredeljaetszja otnoseniem dlinü dugi k radiuszu</u> 'A szög nagyságát az iv hosszának a sugárhoz való viszonya határozza meg'</p>	<p><u>Magnitude of angle is determined by the relation of length of arc to radius</u></p>

Például a 6. sz. mondat szótára így festett:

orosz szavak	angol megfelelőjük		kulcsszámok		
	I	II	I	II	III
<u>velicsina</u>	<u>magnitude</u>	+ + +	+ + +	+ + +	+ +
<u>ugl -</u>	<u>coal</u>	<u>angle</u>	121	+ + +	25
<u>-a</u>	<u>of</u>	+ + +	131	222	25
<u>opredeljaetszja</u>	<u>is determined</u>	+ + +	+ + +	+ + +	+ +
<u>otnoseni-</u>	<u>relation</u>	<u>the relation</u>	151	+ + +	+ +
<u>-em</u>	<u>by</u>	+ + +	131	+ + +	+ +
<u>dlin-</u>	<u>length</u>	+ + +	+ + +	+ + +	+ +
<u>-i^③</u>	<u>of</u>	+ + +	131	+ + +	25
<u>dug-</u>	<u>arc</u>	+ + +	+ + +	+ + +	+ +
<u>-i</u>	<u>of</u>	+ + +	131	+ + +	25
<u>k</u>	<u>to</u>	<u>for</u>	121	+ + +	23
<u>radiusz-</u>	<u>radius</u>	+ + +	+ + +	221	23
<u>-u</u>	<u>to</u>	+ + +	131	+ + +	+ +

③ Az -i és az -ü végződést nem különböztetik meg.

A gép a helyes fordítást az alábbi hat szabály betartásával végezte el:

1. s z a b á l y : Ha az adott szó első kulcsszáma 110, megállapítandó az előző teljes szó harmadik kulcsszáma. Ha ez 21, akkor megváltoztatandó e szavak szórendje (vagyis a 21-es kulcsszámu szónak kell a 110-es kulcsszámu szót követnie); az ellenkező esetben a szórend változatlan marad. Mindkét esetben a 110-es kulcsszámu szó első angol megfelelője veendő.

2. s z a b á l y : Ha az adott szó első kulcsszáma 121, megállapítandó a rákövetkező teljes szónak vagy szótagnak (tőnek vagy toldaléknak) a második kulcsszáma. Ha ez a második kulcsszám 221, akkor a 121-es kulcsszámu szónak az első angol megfelelője veendő; ha a második kulcsszám 222, akkor a második angol megfelelőt kell venni. A szórend változatlan mindkét esetben.

3. s z a b á l y : Ha az adott szó első kulcsszáma 131, az előző teljes szónak vagy szótagnak (tőnek vagy toldaléknak) pedig 23 a harmadik kulcsszáma, akkor a 131-es kulcsszámu szónak a második angol megfelelője veendő, s a szórend változatlan marad. Az ellenkező esetben a 131-es kulcsszámu szó első angol megfelelője veendő, s a szórend megfordítandó.

4. s z a b á l y : Ha az adott szó első kulcsszáma 141, az előző teljes szónak vagy az előző elválasztó szó bármely tagjának (tővének vagy toldalékának) pedig 241 a második kulcsszáma, akkor az adott szónak az első angol megfelelője veendő; ha viszont a második kulcsszám 242, akkor az első angol megfelelő veendő. A szórend mindkét esetben változatlan marad.

5. s z a b á l y : Ha az adott szó első kulcsszáma 151, a rákövetkező teljes szónak vagy bármely tagnak (tőnek vagy toldaléknak) pedig 25 a harmadik kulcsszáma, akkor a 151-es kulcsszámu szónak első angol megfelelője veendő; az ellenkező esetben a második angol megfelelő veendő. A szórend mindkét esetben változatlan marad.

6. s z a b á l y : Ha az adott szó első kulcsszáma + + +, akkor e szónak az első angol megfelelője veendő, és az előző szóval a szórend változatlan marad.

A gép a sorrendnek megfelelően a velicsina szóhoz a hatodik, az ugl- szóhoz a második, az -a-hoz a harmadik, az opredeljaetszja szóhoz a hatodik stb. szabályt alkalmazta.

A fordítógép gyakorlati munkája a következőképpen fest: a gép elején lévő szerkezeten a gépirónó lekopogja az orosz mondatokat latin betűkkel; ezek lyukkártyákra tevődnek át, amelyek meghatározott módon elhelyezett lyukakkal ellátva lépnek be a gépbe, 6-8 másodperc múlva, mint mondtuk, a gép kivezető szerkezete leírja a fordítást.

A szakértők szerint most az a további teendő, hogy elkészítsék az egyes szaktudományok (például az orvostudomány, a kémia, a matematika) szótárait. Ha elegendő tapasztalatot szerezték már, áttérhetnek majd a szépirodalmi művek, a sajtó stb. szövegeinek fordításával kapcsolatos kísérletekre. Ez utóbbi feladat megoldásáig azonban még nagy-nagy nehézségekkel kell megküzdeni: egyebek között a hat szabály helyett, úgy vélik, körülbelül százra lesz szükség. A gép segítségével való fordításnak különleges akadályai a különféle idiómák, sajátos kifejezések. Például a charleyhorse angol szót a gép "Charley (név) ló" kifejezéssel fordítja, holott azt jelenti, hogy "lábgörcs".

Egyelőre tehát csak az első lépéseket tették meg a gépi fordítás felé. Ezek egyrészt azt mutatják, hogy a fordítógép gyakorlatilag megvalósítható, másrészt azokra a nagy nehézségekre is utalnak, amelyek e kérdés további kidolgozásával járnak.

(Voproszű jazükoznanija,
1955/6, 145-8.)

I r o d a l o m :

1. "Referativnűj zszurnal. Matematika" [Tájékoztató lap. Matematika], 1954/10, 75-8.
2. "Journal of the Franklin Institute", 1954/3, vol. 257, 257-60.
3. "Atomes, tous les aspects scientifiques d'un nouvel age", 1954/96, 90-1.
4. "Chemical and Engineering News", 1954/4, vol. 32, 340-1.