

## INFORMÁCIÓS GÉPESÍTÉSI KISÉRLETEK ELEKTRONIKUS SZÁMÍTÓGÉPEL

Vásárhelyi Pál

I.

### A TÉMAFIGYELÉS GÉPESÍTÉSE

Célszerű-e, és ha igen, lehetséges-e hazánkban az elektronikus számítógépek igénybevétele szakirodalmi tájékoztatási célokra? Erre a kérdésre kerestük a választ a külföldi tapasztalatok elemzése és hosszabb magyarországi kísérletsorozat megkezdése útján.

Információvisszakeresés,  
témafigyelés, vagy  
indexkészítés?

Az elektronikus számítógépek dokumentációs felhasználásáról hallva, a tájékoztatási szakemberek általában az információvisszakeresés gépesítésére gondolnak. Ez érthető is, hiszen erről irnak és beszélnek a legtöbbet. Hazai kísérleteink so-

rán mégsem ezzel foglalkozunk először, hanem a témafigyelés gépesítésének és az indexek gépi előállításának magyarországi megvalósítását kívánjuk előkészíteni, s bár az információvisszakeresés alapvető programjainak kidolgozása is folyamatban van, ezzel a kérdéssel behatóbban csak az előző két probléma megoldása után tartjuk célszerűnek foglalkozni.

Ennek oka a következő. Az információvisszakeresés gépesítése akkor válik feltétlenül szükségessé, amikor több százszoros dokumentumból kívánjuk rendkívül nagy sebességgel és igen mély tartalmi feltárás alapján, jól meghatározott, specifikus, vagy éppen igen széleskörű kérdésekre a megfelelő dokumentumokat kikeresni. Nyilvánvaló azonban, hogy a gépi visszakeresés csak akkor valósítható meg, ha a teljes információbázis, tehát a több százszoros dokumentumnak legalább azonosítási száma és valamennyi jellemzőnek ítél't tárgyszava /- sőt lehetőleg teljes címléírása is -/ olyan formában áll rendelkezésre, amilyenben azt a számítógép fel tudja dolgozni. Ma ez a kiinduló gé-

pi bizonylat csak a lyukszalag, vagy lyukkártya lehet.

Az óriási dokumentumtömeg adatainak szalagba vagy kártyába lyukasztása viszont költséges, hosszadalmas és önmagában véve még terméketlen munka, hiszen a tulajdonképpeni tájékoztatást, információvisszakeresést csak hónapok, sőt esetleg évek múlva, vagyis akkor segítheti elő, amikor már a teljes anyag, vagy annak legalább jelentős része feldolgozást nyert. Hazánkban nem célszerű, hogy ilyen hosszadalmas és csak a későbbiek során gyümölcsöző munkát indítsunk be. /Természetesen csak az előkészítés első időszaka ilyen nehéz, mert később a meglévő bázis kiegészítése folyamatosan történik, miközben a gépesített tájékoztatási központ már operatív munkát tud végezni./

A tájékoztatás gépesítésének két másik formája: a gépi témafigyelés és indexkészítés ezzel szemben a gépi bizonylatok biztosításához hosszadalmas és költséges előkészítő munkát nem igényel, tekintettel arra, hogy mindig csak a legfrissebb irodalomra terjed ki, és az azonnali tájékoztatást szolgálja. Ha a szükséges gép és a programok rendelkezésre állnak, akkor már a következő meghatározott periódusban /hónapban vagy negyedéven/ beérkező dokumentumok adatainak szalagba vagy kártyába rögzítése után azonnal konkrét, hasznos és operatív jellegű tájékoztató munka indulhat meg a számítógép felhasználásával. Ez az oka annak, hogy a külföldi cégek, különösen a kisebb vállalatok esetében mindinkább terjed az ún. SDI /Selective Dissemination of Information = Szelektív Információ Terjesztés/ rendszere, és hogy ma már széleskörűen használják a gépi uton előállított gyorstájékoztatókat, indexeket. Ez utóbbiakkal kapcsolatban elegendő, ha csak a kutatóink által nagyra értékelt Chemical Titles, Current Contents vagy orvostudományi területen az Index Medicus című indexkiadványokra utalunk.

A gépi témafigyeléshez, vagy gépi indexkészítéshez kiinduló bizonylatként előállított lyukszalagok vagy lyukkártyák az idők során felhalmozódva, létrehozzák aztán azt a hatalmas információ bázist, amely már a gépi információvisszakeresést is lehetővé és kifizetődővé teszi. Így megtakarítható az inproductív előkészítő időszak.

**Hazai tájékoztatás-  
gépesítési lehetőségeink**

A tájékoztatás gépesítése tehát, amennyiben a témafigyelés, vagy különösképpen az indexkészítés gépesítésére gondolunk, célszerűnek látszik. Kérdés ezután, hogy jelenlegi ill. a közeljövőben kialakuló lehetőségeink

között a gépesítés megoldása lehetséges-e? A válasz e kérdésre ugyan csak kedvező. Nem arról van ugyanis szó, hogy egy tájékoztató központ kizárólag dokumentációs célokra beszeressen egy számítógépet, csupán arról, hogy az országba egyre nagyobb számban beérkező ilyen berendezéseken hetente vagy havonta néhány órát béreljen. A számítógépek bérletének lehetősége az, ami a tájékoztatásgépesítést - legalábbis a témafigyelés és indexkészítés gépesítésének formájában -

már a közeljövőben reálissá teszi. Már most is van az országban olyan számítógép, amely a tájékoztatásgepesítési kísérleti munkák elvégzésére alkalmas és dokumentációs intézmény számára is hozzáférhető: a Számítástechnikai és Ügyvitelszervezési Vállalat GIER típusú számítógépe. Ez év végén, jövő év elején pedig már olyan nagyberendezések is beérkeznek, amelyek a kísérletsorozattal megalapozott operatív jellegű számítógépes tájékoztatást is lehetővé teszik.

A fenti megfontolások alapján kezdte meg a KGM MTTI azt a kísérletsorozatot, melynek célja különböző témafigyelési gépi programok kidolgozása, majd pedig a gépi indexek rendszeres kiadásának előkészítése.

**Az SDI rendszer**

A gépi témafigyelés már említett SDI rendszerét az IBM cég dolgozta ki. A rendszer célja annak megállapítása, hogy az újonnan beérkező dokumentumokról a vállalat vagy intézmény mely munkatársait kell értesíteni oly módon, hogy az érdekelt a számára fontos valamennyi dokumentumról értesítést kapjon, de ne kapjon értesítést semmilyen számára érdektelen anyagról.

Ennek megvalósítása érdekében a gépi témafigyelési szolgálatot igénybevevő szakemberek érdeklődési körét kérdőívek és személyes megbeszélések alapján tárgyszavakkal jellemzik, vagyis meghatározzák az érdekelt profilját, s a profil változását az igénylő visszajelentései alapján állandóan figyelemmel kísérik. Hasonlóan tárgyszavakkal történik a dokumentumok jellemzése is, a témafigyelési értesítések pedig úgy készülnek, hogy a dokumentum tárgyszavakat és a profiltárgyszavakat gépi uton egybevetik és meghatározott arányú megegyezés esetén a dokumentum teljes címleírását tartalmazó értesítést küldenek ki az érdekeltnek.

Az SDI rendszer a tapasztalatok szerint jól bevált, az érdekelteknek megfelelő tájékozottságot biztosít az új irodalommal kapcsolatban, ezért hazai kísérleteink is e rendszer magyarországi alkalmazását kívánják megoldani.

**A profiltárgyszavak meghatározása**

A profilok meghatározását személyes megbeszélések és igényfelmérés útján kezdtük meg. Kísérletképpen különböző szinten dolgozó, automatizálással foglalkozó szakembereket kértünk fel igényeik megadására: akadémiai kutatóintézet, egyetemi tanszék, KGM kutatóintézet és KGM vállalat szintjén dolgozó szakemberektől kaptunk "profilokat".

Az operatív munka során a konkrét kutatási-fejlesztési feladatokból célszerű kiindulni, és a népgazdasági szinten legfontosabb

feladatok megfogalmazását /-természetesen a hozzájuk csatlakozó teljes problémakörrel kiegészítve -/ tekinthetjük profilnak. A legújabb irodalomról szóló értesítéseket e feladatok felelőseinek küldhetjük ki azzal, hogy egyidejűleg felkérjük őket: érdeklődési körük változását, problémáik megoldását, illetve új problémakörök felmerülését közöljék velünk. Ez a "visszacsatolás" alapvető feltétele a gépi témafigyelés eredményes működésének.

**A dokumentumtárgyszavak meghatározása**

A dokumentumtárgyszavak meghatározására több módszer is kínálkozik:

a/ Nagy szakképzettségű dokumentalista a cikk elolvasása alapján saját szavait felhasználva önállóan meghatározhatja azokat a tárgyszavakat, illetve tárgyszó csoportokat, melyeket szakmai ismeretei alapján a dokumentum tartalmára legjellemzőbbnek ítél.

b/ Kis szakképzettséggel rendelkező dokumentalista a dokumentum címében, esetleg rendelkezésre álló referátumában, illetve a dokumentum bekezdéseinek első és utolsó mondataiban aláhúzással kijelölheti a jellemzőnek ítélt tárgyszavakat, vagy szócsoportokat.

c/ A gépi nyelven /lyukszalagon vagy lyukkártyán/ rendelkezésre álló címből és esetleg referátumból a számítógép maga választja ki a tárgyra jellemző tárgyszavakat.

Mi a tárgyszavak meghatározásának először a b/, majd pedig a c/ utját kívánjuk alkalmazni. Az a/ lehetőség megvalósításához megfelelő számú szakember ugyanis nem áll rendelkezésre. Hasonló tendencia érvényesül a gépi témafigyelés külföldi megoldásai esetében is, tekintettel arra, hogy a szakemberhiány világszerte érezhető jelenség.

Érdekes megjegyezni egyébként, hogy a tárgyszavak meghatározásának b/ pontban említett leegyszerűsített, kis szakértelmet igénylő módszerét a teljesen automatizált megoldások közös elemeinek kianalizálása eredményeképpen dolgozták ki. A gépek ugyanis a külföldi kísérletek tapasztalatai szerint elfogadható pontossággal meg tudják állapítani a dokumentumok mondanivalójára jellemző tárgyszavakat, statisztikai és matematikai-logikai alapon oly módon, hogy kiemelik a dokumentumok címében, referátumában, illetve a bekezdések első és utolsó /- s a tapasztalatok szerint általában legnagyobb információ tartalmu -/ mondataiban szereplő lényeges szavakat, s esetleg még a szavak előfordulási gyakoriságát is figyelembe véve állítják össze a tárgyszójegyzéket. A névelőket és más hasonló értéktelen szavakat ennek során természetesen figyelmen kívül hagyják.

A tárgyszavak tárolása  
a gépi memóriában

Mind a profiltárgyszavak, mind a dokumentumtárgyszavak rögzítése két módon oldható meg: direkt és fordított tárolási sorrendben. Az érdekeltek esetében pl. a profiltárgyszó után feltüntethetjük mindazoknak az érde-

kelteknek a nevét, akiknek profiljában az adott profiltárgyszó szerepel és fordítva, az érdekeltek neve után feltüntethetjük teljes profiljuk összes tárgyszavait. Pl.:

tranzisztor

Kis János

Kótai Kálmán

Nagy Jenő

pneumatikus szabályzók

Horváth Kálmán

Nagy József

Sarkadi Károly

.....

.....

illetve:

Kis János

tranzisztor

műszer

táv mérés

Kótai Kálmán

tranzisztor

szabályozás

.....

.....

Hasonló a helyzet a dokumentumok címe és tárgyszavai esetében, ahol ugyancsak megvan mind a direkt, mind az invertált tárolási sorrend lehetősége.

Gyakorlati megfontolásokból kiindulva úgy döntöttünk, hogy a gépi témafigyelés esetében mind az érdekeltek, mind pedig a dokumentumok adatainak rögzítésekor a direkt tárolási sorrendet használjuk. A dokumentumok esetében ez nem igen lehetett vitás, mivel a folyamatosan beérkező dokumentumok feldolgozása nyilvánvalóan úgy a legegyszerűbb, ha a dokumentum címe után azonnal feltüntetjük a jellemző tárgyszavakat, és így rögzítjük az anyagot a lyukszalagba. A gép me-

móriáján ez az anyag egyszerűen "keresztül fut", az összehasonlítások elvégzése után a gép azt "elfelejti", így átrendezésnek nincs értelme. Nem ilyen egyszerű a döntés a profilok rögzítése esetén, melyeket célszerű állandóan a gép memóriájában tárolni. Itt azért döntöttünk ugyancsak a direkt sorrend/név után a tárgyszavak/mellett, mert így a legegyszerűbb a profilváltozások keresztülvétele: a memóriában egy névhez tartozó blokkot egyszerűen törölni kell, és a nevet a helyesbitett új profillal együtt egyetlen blokk-ként ismét be lehet írni a memória egy más helyére. Fordított tárolási sorrend esetén egy profil megváltozásakor esetleg több blokkot is módosítani kellene.

#### A számítógép programozása

A tárolási sorrend eldöntése után arról kell néhány szót szólnunk, hogy hogyan kell utasításainkat a gép számára megfogalmazni, hogyan kell a gépet programozni?

A számítógép számára az utasításokat többféle, egyszerűbb és bonyolultabb "nyelven" adhatjuk meg. Minden gépnek megvan a saját belső utasítás rendszere: "anyanyelven", a gépi kód, és tulajdonképpen csak az ilyen gépi kódban megadott utasítások megértésére és végrehajtására képes. Az utasítások gépi kódban történő közlése azonban rendkívül nehézkes, mert a legapróbb részleteket is külön-külön utasításban kell a számítógép tudtára adni. Egy összeadás programozása pl. lényegében a következő utasításokból tevődik össze:

- vedd ki az A memóriarekesz tartalmát és vidd be a központi számítógépségbe,
- vedd ki a B memóriarekesz tartalmát és vidd be a központi számítógépségbe,
- B értékét a központi számítógépségben add hozzá A értékéhez,
- az eredményt vidd ki a C memóriarekeszbe.

Elképzelhető ezek után, hogy hány ezer lépésből tevődik össze egy bonyolultabb program, pl. a mi esetünkben a profiltárgyszavak és dokumentumtárgyszavak összehasonlítása, vagy rendezése!

A programozók munkájának megkönnyítése érdekében ezért a számítógépekre un. fordítóprogramokat is kidolgoznak, melyek lehetővé teszik, hogy a programozó egy-egy rövid szóban/általában angolul/ megfogalmazott utasítását a gép maga írja át, fordítsa le "anyanyelvére" és bontsa fel a gépi kódznak megfelelő apró részletekre. Ebben az esetben a programozó az összeadási utasítást pl. már a következő formában adhatja meg:

- add össze A és B értéket.

Rendkívül hátrányosnak bizonyult azonban, hogy a különböző cégek gé-

peihez kidolgozott ilyen egyszerűsített programnyelvek különbözőek és így az egyik gépre kidolgozott program a másik gépen nem futtatható. Ennek kiküszöbölése céljából több nemzetközi tanácskozást hívtak már össze, és kidolgoztak néhány olyan programnyelvet, melyekhez a nagy számítógépgyártó cégek ma már általában /- bár nem kötelezően -/ elkészítik a saját géptük "anyanyelvére" átfordító fordítóprogramot, s amely programnyelvek így nemzetközinek tekinthetők. Így elkészült a tudományos problémák megoldásának programozási céljaira az Algol, a kereskedelmi jellegű programozásra a Cobol.

Az Algol elnevezésű nemzetközileg kidolgozott programnyelvet a hazánkban felállított GIER számítógép is ismeri, és a továbbiak során nagyobb számban beérkező ICT számítógépek is megértik. Ezért döntöttünk úgy, hogy programjainkat ezen a nyelven dolgozzuk ki.

Ily módon a programozás viszonylag egyszerű, - de először egy problémát kellett megoldani. Az Algol programnyelv ugyanis, mint mondtuk, elsősorban tudományos számítási feladatok megoldására készült, és így csak számokkal képes műveleteket /pl. összehasonlítást/ végezni, szavakkal nem. Így mindenekelőtt egy olyan programot kellett kidolgozni, mely a szavakat kódszámokkal helyettesíti, hogy a bonyolult témafigyelési programot egyszerű Algol programnyelven adhassuk meg.

A szótárprogram

A szótárprogram két egymástól jól elkülönülő részből tevődik össze: a szótárbeviteli programrészből és a kódoló programrészből. A program első része a szavakból és a hozzájuk rendelt kódszámokból álló "szótár" gépi memóriába történő bevitelére szolgál, míg a második részét akkor használjuk fel, amikor a dokumentumokból kijelölt tárgyszavak kódszámait kívánjuk az előbb bevitt szótár segítségével automatikusan meghatározni és szalagba lyukasztani.

A szótárt a következő felépítésű szalagon visszük be a gépbe:

szó, szám  
szó, szám  
szó, szám  
.  
.  
v,

A beolvasáskor a számítógép mindig csak a szó első 7 betűjét veszi figyelembe, a többit levágja. /l. ábra/ A vessző és a soremelész közötti jeleket, tehát a számokat a lyukszalag első átfutásakor a gép figyelmen kívül hagyja. A szavak memóriába történő bevitele betűnként történik: minden szó számára 7 memóriarekeszt tartottunk fenn. Ha a szó végét jelző vessző /- rövidebb szó esetén -/ már a 7 rekesz kitöltése előtt megjelenik, akkor a fennmaradó rekeszekre a gép nullákat ír. Az egyedülálló v betű jelzi, hogy a szavak beolvasásának végére értünk.

|                |
|----------------|
| positioning,66 |
| pressure,67    |
| program,68     |
| pulse,69       |
| pump,70        |
| radiation,71   |
| radio,72       |
| railroad,73    |
| receiver,74    |
| recording,75   |
| regulation,76  |

1. ábra

Részlet az angol szavakból  
és a hozzájuk rendelt kódszámokból álló "szótárból"

A szavak beolvasása után a lyukszalagot ismét a számítógép olvasóegységébe helyezzük, és olyan utasítást adunk a gépnek, melynek hatására most a szavakat hagyja figyelmen kívül és a számokat viszi be sorra egymás után a memóriába.

A szótárbeviteli programrésszel tehát ily módon egy szó-sorozatot és egy szám-sorozatot vittünk be a gépbe. A kódoló programrész lefuttatásának megkezdésekor olyan lyukszalagot helyezünk az olvasóegységbe, mely a legujabban feldolgozott dokumentumok tárgyszavait tartalmazza eredeti alfabetikus formájukban:

szó,  
szó,  
szó,  
.  
v,  
szó,  
szó,  
.  
v,

Ebben az esetben az egyedülálló v betűk az egyes dokumentumokat jellemző tárgyszósorozatok végét jelzik. A dokumentumok tárgyszócsoportjai a dokumentum-azonosítási számok sorrendjében követik egymást /2. ábra/.

Ez után a számítógép vezérlőpultján elhelyezett írógépen leírjuk /- s így a gép tudtára adjuk -/, hogy hányas számú dokumentumnál kezdjük adott esetben a feldolgozást. /Erre azért van szükség, mert

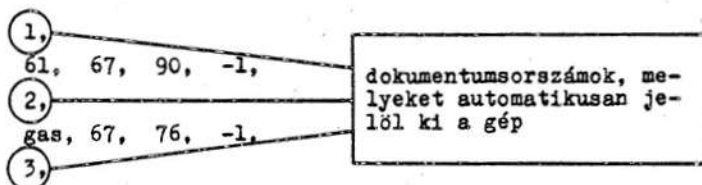


pipeline,  
pressure,  
telemetered,  
v,  
gas,  
pressure,  
regulation,  
v,

2.ábra

Részlet a dokumentumok kijelölt tárgyszavait tartalmazó és a szótárral gépi uton összehasonlítandó tárgyszójegyzékből

a dokumentum-azonosítási számok az egyértelműség biztosítása érdekében a gépi témafigyelés megkezdésétől állandóan emelkedő sorszámok./ A gép ezt a számot rögtön be is lyukasztja a kimenőegységre felhelyezett lyukszalagba, majd az első beolvasott tárgyszót betűről-betűre összehasonlítja a szótárbeviteli programrészt lefuttatása során a gépbe bevitt szó-sorozattal. Ha megegyezést talál, akkor megállapítja, hogy a szótári szó-sorozat hanyadik tagjánál állt meg, és a számsorozaton ugyanennyit lépve előre, megkeresi a szóhoz tartozó kódszámot, s azt ugyancsak belyukasztja a kimenő-szalagba. Ha megegyezést egyáltalán nem talált, akkor a dokumentum tárgyszót a vezérlőpult írógépén kiírja, és a kezelő megállapíthatja, hogy téves lyukasztásról volt-e szó, vagy pedig olyan új tárgyszó fordult elő, amely a szótárban még nem szerepel /3.ábra/.



3.ábra

Részlet abból a bizonylatból, melyet a számítógép a szótárprogram lefuttatása során készít. A 2. ábrán látható tárgyszavak közül az első hármat megtalálta a szótárban, és a megfelelő kódszámmal helyettesítette, a "gas" szó azonban a szótárban még nem szerepelt, ezért eredeti formájában nyomatatta ki. A "pressure" és "regulation" szavakat ugyancsak helyettesítette /az 1. ábrán látható/ kódszámokkal

Szükség esetén a szótár az új szavakkal bármikor kiegészíthető.

A szótárprogram végeredményeképpen olyan lyukszalagot kapunk,

melynek felépítése:

$$D_1, T_{11}, T_{12}, \dots - 1, D_2, T_{21}, T_{22}, \dots$$

ahol:  $D_1, D_2, \dots$  az első, második stb. dokumentum azonosítási száma, amit a gép a feldolgozás kezdetén az írógépen beütött számtól kezdődően egyesével emelkedve rendel az egyes dokumentumokhoz, és  $T_{11}, T_{12}, \dots$  az első dokumentum első, második stb. tárgyszavának kód-száma,  $T_{21}, T_{22}, \dots$  a második dokumentum első, második stb. tárgyszavának kód-száma, és így tovább. Az egyes dokumentumokat a -1 jelzés választja külön.

Az így kapott lyukszalagot használjuk majd fel a tulajdonképpeni gépi témafigyelési program lefuttatása során bemenő adatszalgként.

A gépi témafigyelési programok

A gépi témafigyelési programokat több változatban készítettük el /4.ábra/. A program futtatásának menete minden esetben azzal kezdődik, hogy a számítógépbe beolvassuk először magát a programot lyukszalagról, majd pedig két táblázatot.

Az első táblázat az érdekeltek nevét és pontos címét tartalmazza, valamint az érdekelteknek nyilvántartásba vételük során kiadott sorszámot:

$$\text{név}_1, \text{cím}_1, N_1, \text{név}_2, \text{cím}_2, N_2, \dots$$

ahol:  $N_1, N_2, \dots$  a sorszámok /5.ábra/.

A második táblázat a beérkezett legújabb dokumentumok szerzőjét, címét és bibliográfiai adatait tartalmazza, valamint a dokumentumok /beérkezési sorrendben adott/ sorszámát:

$$\text{szerző}_1, \text{cím}_1, \text{bibliográfiai adat}_1, D_1, \text{szerző}_2, \dots$$

ahol:  $D_1, D_2, \dots$  a dokumentum sorszámok /6.ábra/.

Ezek után megkezdődik az ún. kiindulási tömb beolvasása, mely az érdekeltek sorszáma után kódolt formában megadja profiltárgyszavakat is:

$$N_1, P_{11}, P_{12}, P_{13}, \dots, N_2, P_{21}, P_{22}, \dots$$

ahol:  $N_1, N_2, \dots$  az egyes érdekeltek sorszáma, és

$P_{11}, P_{12}, \dots$  az első érdekelt,  $P_{21}, P_{22}, \dots$  a második érdekelt stb. első, második stb. profiltárgyszavának kódja /7.ábra/.

A profiltárgyszavak kódját ugyancsak meghatározhatjuk a szótár-program segítségével, tekintettel azonban arra, hogy az érdekeltek száma viszonylag kicsi, semmi akadálya sincs annak, hogy az érdekel-

tekkel folytatott megbeszélés után profiltárgyszavaikat már közvetlenül kódolt formában jegyezzük fel egy kinyomtatott kódszám táblázat felhasználásával.

```

begin integer I,J,Jm,k,db,N;
  integer array KT[1:N],S4T[1:280];
  switch S3W:=D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11,D12,
    D13,D14,D15,D16,D17,D18,D19,D20,D21,D22,D23,D24,
    D25,D26,D27,D28,D29,D30,D31,D32,D33,D34,D35,D36,
    D37,D38,D39,D40,D41,D42,D43,D44,D45,D46,D47,D48,
    D49,D50,D51,D52,D53,D54,D55,D56,D57,D58,D59,D60,
    D61,D62,D63,D64,D65;
  switch S1WC:= C1,C2,C3,C4,C5,C6,C7,C8,C9,C10,C11,C12,C13,
    C14,C15,C16,C17,C18,C19,C20;
  N:=typein;
  input(KT);
L1:typechar; comment adatszalatot berakni;
  I:=1; J:=0;
L2: input(S4T[I]);
  if S4T[I]=-2 then goto KI;
  if S4T[I]=-1 then
  begin
  J:=J+1; if J=25 then goto KI;
  I:=I+1; goto L2;
  end;
  I:=I+1; goto L2;
  
```

4. ábra

Részlet a gépi témafigyelés programjából

Ezután a szótárprogrammal előállított adatszalatagon szereplő dokumentum-tárgyszavak és a profil-tárgyszavak összehasonlítása következik. Az összehasonlítás menetének különböző változatait próbáltuk ki:

1. Első változat:

A gép először kikeresi a második táblázatból az első dokumentum szerzőjét, címét és bibliográfiai adatait, és a gép kimenőegységét képező írógépen kiírja, vagy későbbi kiíratás céljából lyukszalagba lyukasztja azt. Majd a dokumentum minden egyes tárgyszavát sorra összehasonlítja az összes profiltárgyszavakkal, vagyis:

T<sub>11</sub>-et P<sub>11</sub>, P<sub>12</sub>, ... P<sub>21</sub>, P<sub>22</sub>, ... stb.-vel, majd  
T<sub>12</sub>-t ugyancsak P<sub>11</sub>, P<sub>12</sub>, ... P<sub>21</sub>, P<sub>22</sub>, ... stb.-vel.

```
C3: outtext({<Grandl Jozsefne, Bp.VIII.Prater u.39.
});
    goto V;

C4: outtext({<Laki Jozsef, Bp.XIX.Ady E.u.35.
});
    goto V;

C5: outtext({<Szeker Istvan, Bp.I.Logodi u.3.
});
    goto V;

C6: outtext({<Tiborc Erno, Bp.VI.Sziv u.9.
});
    goto V;

C7: outtext({<Tibor Laszlo, Bp.IV.Hunfal L.u.10.
});

    goto V;
```

5.ábra

Részlet az érdekeltek cimleírását tartalmazó táblázatból

Ha megegyezést talál a dokumentum már kiirt /vagy szalagba lyukasztott/ adatai alá /illetve után/, feltünteti annak az érdekeltek nevét és címét, akinek profiljában a megegyezést találta. Ha tehát pl. T<sub>11</sub> megegyezett P<sub>21</sub>-el, akkor N<sub>2</sub>-höz az első táblázatban megkeresi a megfelelő nevét és címet, és azt adja ki.

Ezen alap-program végterméke egy olyan jegyzék, melyben minden új dokumentum adatai után ott szerepel azoknak az érdekelteknek neve

és címe, akiket az adott dokumentumról értesíteni kell.

|      |                                                                                                                                                                                        |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| D34: | outtext({<.....HAYWOOD,H., SOSIN,B.M.<br>Modulált jelek eloallitasa nagyfrekvencias adok szamara.<br>Point to Point Telecommunication, 8.k. 3.sz. 1964.jun. p                          |
|      | <u>goto</u> V;                                                                                                                                                                         |
| D35: | outtext({<.....GAREEV,F.A., POPOV,V.A.<br>Dioda nélküli hidkapcsolású logikai alaparamkorok.<br>Izvesztija vuszsih ucsebmuh zavedenij.<br>Radiofizika, 7.k. 2.sz. 1964. p. 352-357.}); |
|      | <u>goto</u> V;                                                                                                                                                                         |
| D36: | outtext({<.....COATES,C.L., LEWIS,P.M.<br>DOMUT kuszobertekkapukból épített számítógép.<br>IEEE Transactions on electronic computers,<br>13.k. 3.sz. 1964.jun. p. 240-247.});          |
|      | <u>goto</u> V;                                                                                                                                                                         |
| D37: | outtext({<.....Pneumatikus logikai egységek.<br>Measurement and Control, 3.k. 7.sz. 1964.jul. p. 271-272                                                                               |
|      | <u>goto</u> V;                                                                                                                                                                         |

6. ábra

Részlet a dokumentumok címlírását tartalmazó táblázatból

|                               |
|-------------------------------|
| 1, 99,127,145,198,201,404,-1, |
| 2, 95,121,143,156,176,300,-1, |
| 3, 31,43,56,156,301,-1,       |

7. ábra

Részlet a "kiindulási tömbből"

Mivel azonban a legegyszerűbb alap-programról van szó, ha egy dokumentum tárgyszavai közül 2, 3, stb. egyezett meg valamely érdekelt 2, 3, stb. profiltárgyszavával, akkor a dokumentum adatai után az illető érdekelt neve és címe két-három alkalommal is szerepel.

Ezt küszöböli ki a

2. Második változat:

Ebben az esetben a gépnek azt az utasítást adtuk, hogy ha valamely dokumentum és valamely profil tárgyszavai között megegyezést talál, írja ki az érintett érdekelt nevét és címét, de ha ugyanazon érdekelt esetében további megegyezést is talál, akkor már ne írja ki ismét a nevet és címet, csupán számolja meg, hogy hány volt a megegyező tárgyszavak száma, s ezt a számot mint a megegyezés "súlyát" jelző számot, írja ki az érdekelt neve és címe után. Pl.:

J. Schwarz:

Új tranzisztoros mérőműszer

Funktechnik, 13.k. 1.sz. 1965.január p.15-21.

Kiss József, Bp. V. Guszev u. 31. /2/

Úveges György, Bp. II. Pény u. 1. /5/

.....

Az első ill. a tökéletesített második változatot olyan esetekben célszerű felhasználni, amikor a dokumentumokról referátum-kártyák állnak pl. rendelkezésre, s az a kérdés, hogy egy dokumentum referálókártyáit mely érdekelteknek kell megküldeni. Ilyenkor a referálókártyák nyilván dokumentumonként összegyűjtve állnak rendelkezésre és így a csoportosítás helyesen a gép esetében is a dokumentum adataiból indul ki.

3. Harmadik változat:

Ez az előző kettőhöz képest pontosan fordított uton jár. Ebben az esetben ugyanis a gép először az első táblázatból írja ki az érdekelt nevét és címét, majd az összehasonlítási műveleteket a következő sorrendben végzi el:

- a/ az első érdekelt első profil-targyszavát összehasonlítja az első dokumentum minden tárgyszavával,
- b/ az első érdekelt második, harmadik, .... utolsó profiltargyszavát összehasonlítja az első dokumentum minden tárgyszavával,
- c/ az első érdekelt első, második, harmadik stb. profiltargyszavát összehasonlítja a második dokumentum minden tárgyszavával,
- d/ így folytatja az összehasonlítást mindaddig, míg az első érdekelt minden profiltargyszavát össze nem hasonlította valamennyi dokumentum valamennyi tárgyszavával.

Ha megegyezést talál, akkor az érdekelt neve és címe alá ki-

írja a megfelelő dokumentum adatait, majd a második változatban már említett számlálási művelethez hasonlóan jár el, és feltünteti a dokumentum adatai után, hogy hány tárgyszó egyezett meg az adott érdekelt profiltárgyszavaival.

Pl.

Kiss József, Bp. V. Guszev u.31.

J.Schwarz:

Új tranzisztoros mérőműszerek

Funktechnik, 13.k. 1.sz. 1965.jan.p. 15-21.

F. Joó

Tranzisztorizált frekvencia-mérő

.....

- e/ A dokumentumsor végére érve, a számítógép az első táblázatból kiírja a következő érdekelt nevét és címét, és hasonlóképpen elvégzi az összehasonlítást valamennyi dokumentummal.
- f/ Az összehasonlítást mindaddig folytatja, míg valamennyi érdekelt esetében meg nem vizsgálta az összes dokumentumokat.

A munka eredménye ebben az esetben tehát olyan jegyzék, mely minden érdekelt neve és címe után megadja az őt érdeklő dokumentumok adatait, valamint az egyes dokumentumok súlyát. /vagyis azt a számot, amely megmutatja, hogy hány tárgyszó egyezett meg a dokumentum és a profil között./

Az ilyen típusú jegyzék olyan esetekben használható jól fel, amikor nem állnak rendelkezésre olyan kartonok, amelyeket az érdekelt bővebb tájékoztatása céljából ki lehetne küldeni. Ilyenkor ugyanis a számítógép által az érdekeltek neve után összeállított címjegyzék /irodalom-jegyzék/ közvetlenül postázható, manuális munkára azon kívül, hogy a gépből folyamatosan kijövő papirtekercset szét kell vágni és borítékokba kell tenni, szükség nincs.

#### 4. Negyedik változat:

Az előző változat további finomítása olyan értelemben, hogy a gépnek előre megadjuk, milyen súlyozást kívánt az illető érdekelt, és a gép az érdekelt neve és címe alá csak akkor írja ki valamely dokumentum adatait, ha annak súlya az előírt értékkel megegyezik, vagy annál nagyobb.

Az ily módon elkészített jegyzék az előzőnél szelektivebb, biztosítja, hogy az érdekelt olyan irodalom-mennyiségről kapjon tájékoztatást, amely kívánságának megfelelően se nem túl sok, se nem túl kevés.

5. Ötödik változat:

Ez a változat /- mely még kidolgozásra vár -/ lehetővé teszi, hogy mind a dokumentum - mind a profiltárgyszavak egymáshoz való kapcsolatát is feltüntessük, és így az esetleges félreértéseket kiküszöböljük. Ha ugyanis pl. egy cikk tranzisztorok mérésével és a megfelelő műszerek gyártásával foglalkozik, el kell kerülni a tranzisztor és gyártás, illetve a műszer és mérés tárgyszavak kapcsolódását.

További teendők

Összefoglalva az eddig mondottakat, megállapíthatjuk, hogy a gépi témafigyelés hazai megvalósítására irányuló kísérletek eddig eredményesek voltak. A kísérleteket az év folyamán tovább folytatjuk, s annak megvalósítására törekszünk, hogy az idegennyelvű anyagból kiindulva, a gépi témafigyelést egyben a lehető legegyszerűbb gépi fordítási /pontosabban inkább gépi szótározási/ programmal kiegészítve biztosítsuk a magyar nyelvű témafigyelést. E munka során egyben kialakulnak azok a gépi bizonylatok, melyek az információvisszakeresési kísérletekhez biztosíthatják majd a szükséges kiindulási anyagot. Ezzel párhuzamosan meg kívánjuk kezdeni a gépi indexkészítési munkával kapcsolatos kísérletsorozatot is.



VÁSÁRHELYI, P.: Experiments on the mechanization of information by using electronic computers

Of all the possibilities of adopting electronic computers for professional literature information purposes, it was selected as a target objective, above all, to realize the mechanization of current literature search in this country, with regard to the fact that this field is the one to require minimum preparatory work. In course of the experiments, first the so-called vocabulary program was completed, the task of which was to code the subject headings listed alphabetically. This was followed by experiments making use of various alternatives in the field of the very mechanization of the current literature search. The paper explains in detail the sequence of the mechanized process, and its possibilities existing in this country.

.□.



ВАШАРХЕЙ, П.: Испытания по механизации информационной деятельности при помощи электронно-вычислительной машины.

Из возможностей применения электронных-вычислительных машин для информации о спецлитературе, Институт научно-технической информации Министерства металлургии и машиностроения выбрала прежде всего цель осуществления механизации наблюдений за темами в нашей стране, в виду того, что эта работа связана с минимальным расходом подготовок. В процессе исследований сначала была подготовлена так называемая словарная программа, заключающаяся в кодировке дескрипторов в алфавитном порядке. После этого, испытывались различные варианты механизации за темами. Статья автора подробно рассматривает ход механизации и имеющиеся в нашей стране возможности.



VÁSÁRHELYI, P.: Versuche zur Mechanisierung der Informations-tätigkeit mit einem elektronischen Rechenautomat

Unter den Möglichkeiten der Verwendung von elektronischen Rechenautomaten für die fachliterarische Informationstätigkeit wurde in Ungarn als erstes die Mechanisierung des Themenbeobachtungsdienstes angestrebt, da dies die wenigsten Vorbereitungsarbeiten erfordert. Im Laufe der Versuche wurde vorerst das sogenannte Wörterbuchprogramm, das zum Kodieren der alphabetisch angegebenen Sachwörter dient, verfertigt. Sodann wurden Versuche mit verschiedenen Variationen zur Mechanisierung des Themenbeobachtungsdienstes selbst, durchgeführt. Ausführlich wird der Gang der mechanisierten Arbeit mit ihren in Ungarn bestehenden Möglichkeiten beschrieben.

