

Rakov, B. M. - Cserenin, V. P.:

SZOVJET INFORMÁCIÓKERESŐ-GÉPEK

A tudomány és a technika fejlődése nyomán az ismeretek különböző ágaiban a rendelkezésre álló adatok mennyisége szüntelenül növekszik. Ezért egy meghatározott problémára vonatkozó irodalmi adatok kutatása mind nehezebbé válik, s az erre a célra használt dokumentációs eszközök /könyvtári katalógusok, nyomtatott bibliográfiák, stb./ mind kevésbé felelnek meg.

A probléma megoldására szerkesztett és különböző tanulmányokban/1/+ ismertetett mechanikus eljárások főként a lyukkártyák és a mikrofilmek alkalmazásán alapulnak. Ez a módszer azonban, mely szűkebb terület adatainak keresésére sikerrel használható, nagy adattömegek esetén többé nem felel meg a célnak, mivel nehézséggel jár olyan jelzésrendszernek /gépi, tájékoztató nyelvnek/ megalkotása, mely az adatok mechanikus szelektálását lehetővé teszi.

Nagyszámú adatnak mechanikus szelektálása két alapvető problémát vet fel: az egyik a tájékoztató nyelv kidolgozása, a másik különleges technika megalkotása.

A fő nehézség /2, 3/ az adatok osztályozására /helyesebben "nyilvántartására" - Szerk./ szolgáló jelzésrendszer /kódrendszer/ szerkesztésében rejlik, mivel az - többek között - a terminológia egységesítését, a szakkifejezéseknek az osztályozás céljához való hozzáhangolását, az index-szerkesztést, stb. is felöleli.

Az információkereső-rendszer több tényezőtől függ: ama munka jellegétől, amelynek érdekében fel akarják használni, továbbá a rendelkezésre álló információ anyag mennyiségétől, az információkérésektől, az információ szolgáltat szervezetétől és működésétől stb. Ha szűkebb ismeretkörrel van esős, egyszerű eszközök, mint pl. peremlyukasztású + a zárójelben lévő számok a cikk végén közölt bibliográfiára utalnak.

kártyák, vagy egymásfőlé helyezhető lyukkártyák /szuperpozíciós, ~~etn~~-zeti lyukkártyák - Szerk./ is megfelelnek. De vannak bonyolultabb eljárások is: a lyukkártyákkal működő adatfeldolgozó gépek /pl. az IBM 101-es gép, a Samain-féle lyukkártyaszелеktor, a Luhn-féle gép/, vagy a mikrofilmekkel dolgozó gépek /pl. a Samain-féle filmszelektor, a Filmorex-szelektor, a Bush-féle mikrofilm-szelektor/.

V. P. Cserenin tanulmánya /3/ rámutat a külön e célra készült berendezésnek, azaz a speciális információkereső-gépnek ama előnyeire, amelyek egy "gépi /tájékoztató/ nyelv"-nek megalkotásakor, szerkezetének tanulmányozásakor s a /fogalomkapcsolások létesítésére, jelzésrendszer megszerkesztésére, stb. vonatkozó/ kísérletek lefolytatásakor jelentkeznek. E gépeknek az a rendeltetésük, hogy egy tárgykör jellemzőinek /karakterisztikumainak - Szerk./ összességéből az információ céljainak megfelelő egyedi tételeket kiválogassák. A gépek leglényegesebb tulajdonságának nem abban kell megnyilvánulnia, hogy hatalmas adattömeget nagy sebességgel tudjanak átválogatni, hanem, hogy nagyszámu, egymástól különböző jelzet szerinti keresésre legyenek alkalmazhatók. Alkalmasaknak kell lenniök tehát egy tárgy különböző jellemzői közötti logikai kapcsolatok felismerésére és nagyszámu jellemző egyidejű tekintetbevételére.

A Szovjetunió Tudományos Akadémiájának Tudományos Tájékoztató Intézete 1954-ben e célra kísérleti információkereső-gépet /EIM/ szerkesztett /4/. A gép ugyanazon alapelv szerint működik, mint a standard adatfeldolgozó gépek, ugyanugy 80 oszlopos lyukkártyákat használnak hozzá. Bár szerkezete egyszerű, mégis minden követelménynek teljesen megfelel, s a kívánt üzemeltetési hajlékonysággal /sokoldalúsággal - Szerk./ is rendelkezik. Olyan intézményeknél, vállalatoknál is alkalmazható tehát, amelyeknek már van adatfeldolgozó gépük.

A lyukkártyák jelzésrendszere és kártyaterve

Minden információs tétel /dokumentum/ különböző karakterisztikumok alá sorolható. Ezeket a jellemzőket az információkereső-gépek számára a "gép nyelvére" kell áttenni és a tárgy feldolgozásának első lépéseként jelzetek /kódjelek/ formájában kell megállapítani. Az egyes információs tételekre /dokumentumokra/ vonatkozó jelzetek összességét

aztán egy 80 oszlopos lyukkártyára lyukkombinációkkal jelölik be. Ebből áll a tárgy feldolgozásának második lépése. Így történik ez az EIM-gépnél is.

Számos olyan adatbejelölő eljárás van, amelynél a jelzetek leképezése lyukasztások által történik. A tapasztalat nyomán /3/ a különböző jelzésrendszerek a második lépést jelentő kód-leképezés természete és kódképeknek a kártyán /a jelen esetben a lyukkártyán lyukak útján/ való elhelyezése szerint az alábbi csoportokba sorolhatók:

1/ Jelzésrendszer, amelynél a jelzetek a kártya megfelelő jelhelyeihez vannak kötve /helyhez kötött jelzetek, fix elhelyezésű jelzetek/. Ezeknél a karakterisztikumok leképezését az alkalmazott jelzet, valamint ennek a kártya megfelelő helyén történő bejelölése együttesen adja.

2/ Jelzésrendszer, amelynél a jelzetek nincsenek a kártya kiválasztott jelhelyeihez kötve /szabadelhelyezésű jelzetek, helyhez nem kötött jelzetek - Szerk./. Itt is megvan minden karakterisztikumnak a saját jelzete, s minden jelzet felvételére a kártyán külön jelmezőszelvény szolgál, e jelmezőszelvények elhelyezése azonban kötetlen.

3/ A jelzetek egymásra helyezésén alapuló rendszerek /szuperpozíciós rendszerek - Szerk./. Ebben az esetben a különböző karakterisztikumok jelzeteit, amelyek szimbolumaikat, illetően különböznek egymástól, a kártya ugyanazon jelmezőjébe, egymásra reáhelyezetten jelölik be. Az ilyen rendszereknél fennáll a fiktív jelzetek képződésének lehetősége.

4/ A közvetlen leképezésű jelzésrendszerek /direkt-jelzésrendszerek - Szerk./, amelyek lényegében véve az 1/ alatt említett eljárásan alapulnak. Itt azonban a karakterisztikumok jelzetei abban különböznek egymástól, hogy mindegyikük számára külön hely van felvéve a kártyán, ahol csupán kétféle bejelölés eszközölhető; ezek közül az egyik a szóbanforgó karakterisztikum meglétét, a másik ennek hiányát jelenti.

A helyhez kötött jelzetek az egy tárgykörhöz tartozó, egymást kölcsönösen kizáró jellemzők leképezésére használatosak. Minden tárgykörnek megvan a kártyán a saját jelmezőszelvénye, amelybe e tárgykör bármelyik egyedi jellemzője saját jelzetével jelölhető be. A "kölcsönösen kizáró" jelzöt azért használtuk, mert a dokumentumokban valamely jellemző-csoportnak több egymást kölcsönösen kizáró jellemzője is előfordulhat.

A direkt-jelzettek alkalmazása akkor hasznos, ha a dokumentumban egy viszonylag kisebb mennyiségű jellemző-tömegben viszonylag sok jellemző fordul elő. Így a 80 oszlopos kártya használata maximálisan 960 darab karakterisztikumot felölelő jellemző-tömeg számára elegendő. Ezek közül az egyes jellemzőket a kártya megfelelő jelhelyére tett egyetlen lyukasztás képezi le.

A legtöbb célra alkalmazható és a helyet jól kihasználó jelzésrendszer az, amely nincs a kártya kiválasztott jelhelyeihez kötve /helyhez nem kötött jelzettek/, mivel ennél a rendszernél csak a dokumentumban tényleg fellelhető jellemzők számára kell a kártyán helyet biztosítanunk. A karakterisztikumok leképezésére szolgáló jelmezőszelvények minden előzetes rendezés nélkül /tetszés szerint/ következhetnek egymás után; - úgy kell azonban elhelyezve lenniök, hogy a gép folyamatosan érzékelhesse le őket.

A szuperpozíciós jelzésrendszereknél a jelmezőszelvénybe egyetlen jellemző helyett több jelölhető be, és így a kártyára felvihető adatok mennyisége jelentősen növelhető. Viszont rontja a rendszer hatásosságát az a feltétlenül tekintetbe veendő körülmény, hogy a keresett tételek mellett fölös, nem keresett tételeket is kiválasztanak.

Az EIM egyaránt dolgozik az említett kódrendszerek bármelyikével, vagy azok bármilyen kombinációjával, mint: szuperpozíciós + helyhez nem kötött; szuperpozíciós + helyhez kötött; helyhez kötött + helyhez nem kötött rendszerekkel. Ilyen esetben a kártya zónákra van osztva, s mindegyik zónába a vonatkozó rendszer jelzete kerül. Végeredményben tehát minden kártya négy zónára osztható: a helyhez kötött jelzettek, a direkt-jelzettek, a helyhez nem kötött jelzettek, a szuperponált jelzettek zónáira.

A jellemzőknek az egyes zónákban történő leképezéséhez /ha a szuperponált, a helyhez kötött és a helyhez nem kötött jelzettek használjuk/ kétféle jelzettípus /kódtípus/ használható: a/ az egyiknél a zóna összes kód-konfigurációi azonos számú jel-elemből állanak /szelktív- vagy ekvivalens jelzettek/; b/ a másiknál viszont a különböző konfigurációk egymástól eltérő számú jel-elemet tartalmazhatnak /ordinális vagy kettős jelzettek/.

Az EIM lyukkártyáinak jelzettelésére szelektív-kódok szolgálnak, ami

jelentősen egyszerűbb összehasonlító áramkörök alkalmazását tette lehetővé.

Lássuk mármost az EIM géphez használatos lyukkártyák kártyatervét. A kártya négy /I-IV/ zónára van osztva, amelyekbe már be vannak jelölve a vonatkozó jelzések.

Az I.zóna, mely helyhez kötött jelzetek felvételére szolgál, 10 jelmezőszelvényt / +1 tartalékszelvényt/ ölel fel. Ezek mindegyike több, egymást kölcsönösen kizáró jellemző egyikének jelzetét tartalmazza. Az 1 és 2 jelmezőszelvénybe a lyukkártyagépek szokásos bejelölésével leképezett numerikus adatok kerülnek. A 3-10 szelvények különleges - e célra szerkesztett - jelzésrendszerekkel leképezett numerikus és alfabetikus adatokat vesznek fel. Az I.zóna elrendezése tehát olyan, hogy az adatoknak a kártyán történő elhelyezésében igen nagyfokú rugalmasságot tud biztosítani.

A II.zóna a közvetlen jelzetű jellemzők bejelölésére szolgál. $8 \times 12 = 96$ jelhelyet tartalmaz, ami azt jelenti, hogy egyidejűleg maximálisan 96 jellemző leképezését tudja felvenni.

A III.zóna a szuperpozíciós jelzetekkel leképezett jellemzők zónája. Az egyes oszlopokba különböző - akár betű-, akár számjegyekből álló - kódjelek vihetők be, és a jellemző leképezésére a zóna valamennyi oszlopa igénybevehető. Ez az eljárás például tulajdonneveknek - a lyukkártyák szokásos alfabetikus jelölésmódjával való - bejelölésére használható; ilyenkor a betűket ugyanazon oszlopba tett két-két lyukasztás képezi le.

A IV.zóna a helyhez nem kötött jelzeteké. Egy-egy jellemző felvételére a kártya egy sora, illetve sorának a 41-80 oszlopok közötti része szolgál. A zónában 12 különböző jellemző számára van hely. A jellemzők jelzetei akár betű-, akár számjegyekből állhatnak. Ha például az orosz nyelv 32 betűt tartalmazó alfabetumát vesszük alapul, akkor egy-egy betű leképezéséhez a kártya egy sorának 7 jelhelye szükséges, amelyek közül 3 lyukasztott, 4 lyukasztás nélküli lesz.

Amennyiben a dokumentációs munka jellege megkívánja, mind a kódrendszerek, mind a kártyaterv megváltoztatható. Az elmondottak világosan mutatják, hogy az EIM-hez a különböző típusú jelzésrendszerek egyaránt jól használhatók.

A jelzetek bejelölésére szolgáló gép

Az információs adatoknak a kártyára való felvitelére különleges kódoló lyukasztógép készült. Ez tulajdonképpen a szabványos 80 oszlopos kártyalyukasztógép, amelynek vezetékekben /- ez különleges kódoló áramkör elektronikus elemekkel -/ megfelelő változtatásokat és pótlásokat eszközöltek, hogy a jellemzők kódjainak leképezését a fenti elrendezés szerint vihesse rá a kártyára.

A lyukasztógép olymódon működik, hogy amikor a kártya egy része áthalad rajta, olyan kódrendszerben végzi a bejelölést, amint ama kártyarészen levő zónában éppen szükséges. A bejelölést végrehajtó áramkörök a kártya tovahaladtával más kódrendszerre kapcsolnak át. Az átkapcsolás a kártyaterv szerint beállítható különleges átváltószerkezet segítségével automatikusan történik.

Amikor a helyhez nem kötött jelzetek lyukasztására kerül sor, a gép a lyukasztásokat automatikusan sormenti irányban végzi, s a kártyát annyiszor futtatja a lyukasztóberendezésen keresztül, ahány be-lyukasztandó jelzet van. /A lyukasztógép ugyanis általában az oszlopok irányában végzi a bejelölést. - Szerk./

A számjegyekkel leképezett jellemzők numerikus billentyűzet, a betűjegyekkel leképezettek alfabetikus billentyűzet segítségével lyukasztathatók be.

A jellemzők jelzetei azonban más módon is rávihetők a kártyára. Mégpedig egy kártyamásológép segítségével. Ebben az esetben minden jellemzőről egy külön kártya, az alapkártya készül. /A másológép erről másolja át a vonatkozó jellemzőt leképező lyukasztásokat a dokumentum információs kártyájába. - Szerk./

A lyukasztások helyességének ellenőrzésére minden kártyáról másodpéldány készíthető, melyet azután az első példánnyal össze kell hasonlítani. De megoldható ez az ellenőrzés másképpen is, ti. egyeztetőgéppel, mely lényegében véve a lyukasztógéppel azonos szerkezetű berendezés.

A kísérleti információkereső-gép /EIM/

Az EIM szerkesztése ugyanolyan elvek alapján történt, mint a C-80-1 lyukkártya-feldolgozógépe; az áramkörökben azonban -a szükségessé sokoldalúság érdekében - számos változtatást és pótlást hajtottak végre.

Az EIM gép adatkereső szerkezete a kártya 80 oszlopáról - amíg a kártya átfut a szerkezeten - egyidejűleg érzékeli le a kijelölt adatokat. A gép maximális szelektáló sebessége másodpercenként 7 kártya.

A kívánt adatok kiválogatására külön e célra szerkesztett elektromos áramkörök szolgálnak, továbbá egy olyan szerkezet, mely a feltett kérdéseket a szelektálóba juttatja. A gépnek kétféle szelektáló szerkezete van: az egyik a szokásos kivitelű kapcsolótábla, a másik olyan vezérlő tábla, melyen a kártya minden jelhelye számára külön kapcsolóelem van.

A kívánt adatokat kereső áramkör két részből áll. Az első rész összehasonlító munkát végez: összehasonlítja a kártyába bejelölt jellemzők jegyeit a kérdés jellemzőinek jegyeivel. A második rész ezen összehasonlítás eredményét regisztrálja s ennek következményeként a kártyát a gép gyűjtőrekeszeinek valamelyikébe irányítja. Az összehasonlító szerkezetben sem elektroncsövek, sem relék nincsenek, olcsó és üzembiztosan működő alkatrészekből van építve. Viszont az összehasonlítás eredményét regisztráló szerkezetben már számos elektronikus regisztráló elemet és relét találunk. A helyhez nem kötött jelzetek összehasonlítási eredményeinek regisztrálása céljából annyi regisztrálóelem beépítése szükséges, ahány jellemző jegye jelölhető be a kártyába. Más típusú kódok használata esetén egyetlen regisztráló-elem elegendő.

Az EIM úgy végzi az adatkikeresést, hogy - amíg a kártya a gépen egyszer végigfut - a bejelölt adatok jegyeit /lyukasztásait/ összehasonlítja a feltett kérdés jegyeivel. E célból az EIM áramkörei úgy vannak szerkesztve, hogy a gép a kiszelektálást a kártyába bejelölt jellemzők és a kérdés jellemzői közötti logikai kapcsolat létrejötte alapján végzi. További logikai kapcsolatok létesítésével a szelektálás területe még növelhető. Így például szelektálni lehet A és B, plusz C és D szerint. Föltétve, hogy A, B, C, D 1-el vagy 0-val egyenlő, aszerint, hogy a vonatkozó jellemzők a szelektált tétel jegyei között fellelhetőek-e, vagy sem, továbbá a logikai kizárás szimbolumát használva, az az összefüggés F függvényeként így fejezhető ki: $F = A \cdot B / C + D$. E függvény értéke 1, ha a vonatkozó jelzet szelektálódik, illetve 0, ha nem szelektálódik. A gép úgy is beállítható, hogy más függvényekkel reprezentált fogalomkapcsolások regisztrálására is alkalmas legyen, például

dául a kölcsönös kizárás jegyének használatával, stb.

Az EIM elektromos áramköréit úgy szerkesztették meg, hogy az adatszelektáló szerkezetet akár egy kérdés nagyszámu jellemzője szerinti keresésre, akár több kérdés egyidejű megválaszolására könnyen lehessen bekapcsolni.

Az átválogatásra kerülő kártyákat megfelelő mennyiségenként a gép táprekeszébe kell helyezni. Innen a kártyák a leérezékelő /letapogató/ berendezésbe jutnak, mely azokat az összehasonlítás után és annak eredményeként, az I. vagy II.rekeszek valamelyikébe irányítja. A kijelölt adatokat tartalmazó kártyák az I.rekeszben, a többiek a II.-ben gyűlnek.

A feladott kérdés jellemzőit vagy a kapcsolótáblán vagy a szelektoron kell beállítani. A kapcsolótáblán való beállításnál még a szelektálás módja is jelölve van.

Az automatikus adatkereső-berendezés regisztráló elemei külön szerelékklapoken vannak elhelyezve. Ugyanezen van az összehasonlító berendezés is, mely a kérdés jelzeteit a kártya jelzeteivel veti össze.

Leíró-szerkezet

Külön író /gépelő-/ szerkezet szolgál a bibliográfiai adatok kiírására a kiszelektált kártyákról. E bibliográfiai adatok felölelik pl. a katalógusszámot, a dokumentum számát vagy rövidített címét, stb. Mivel az adatok mind rövidek, gyorsan kiírhatók. Ezért tabelláló gép helyett olyan leírószerkezetet lehet alkalmazni, amely nagyjából a kártyaellenőrző gép elvei szerint van szerkesztve, - s szolenoiddal ellátott írógép-ből áll, melyet elektromosan kapcsolnak az adatregisztráló szerkezethez. A leírandó lyukkártyákat a regisztráló szerkezet táplálórekeszébe kell helyezni, amely leérezékeli őket. Ennek során a keresett adatok elektromos impulzusokká alakulnak át, amelyek aztán átfutnak az írógép szolenoidján s ennek következtében az írógép automatikusan betű, illetve számjegyek formájában írja le a keresett választ.

IRODALOM:

/1/ A kérdés egész anyagát felölelő bibliográfia tekintetében, az 1951.évig bezárólag, lásd R. S. Casey és J. W. Perry: Punched cards /Lyukkártyák/, New York, 1951.

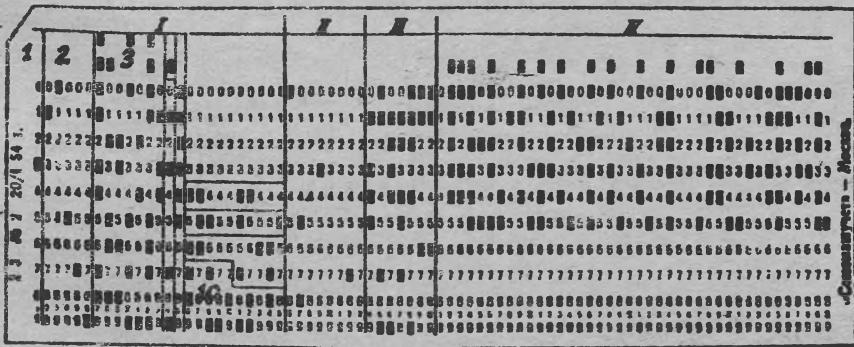
- /2/ J. W. Perry, Allen Kent és M. M. Berry: Machine literature searching /Irodalomkutatás gépi eszközökkel/, New York - London, Interscience Publishers, 1956.
- /3/ V. P. Cserenin: Nekotorüs problémü dokumentacii i mehanizacija informacionnüh poizskov. /A dokumentáció egyes kérdései és az információkeresés gépesítése./ Moszkva, 1955.
- /4/ B. M. Rakov és V. P. Cserenin: Ékszperimental'naja informacionnaja masina Insztituta naucsnoj informacii A.N. SzSzsZR. /A Szovjetunió Tudományos Akadémiája Tudományos Tájékoztató Intézetének Kísérleti információkereső-gépe./ Moszkva, 1955.

.....

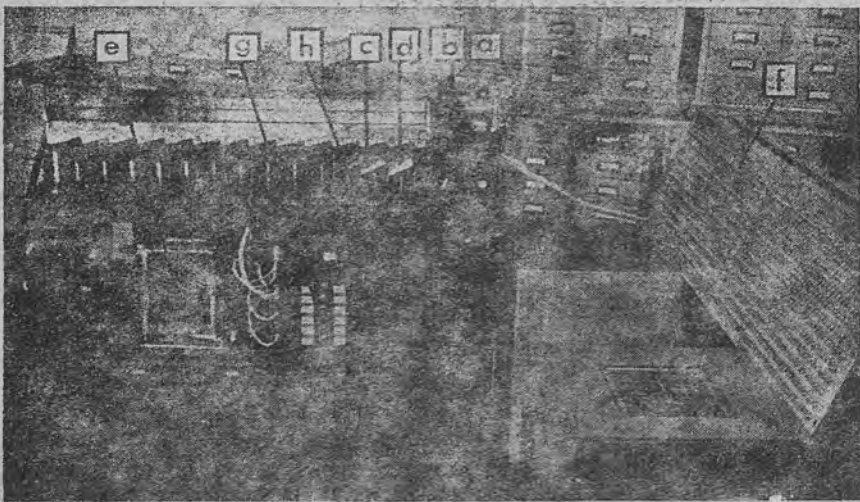
Eredeti cím: Machines for retrieving information in the U.S.S.R.

Megjelent: Unesco Bulletin for Libraries, 1957. 8-9.sz.
192-197.lap.

OSzK.ford.sz.: 1856.



1.sz. ábra
Az EIM-gépnél használt lyukkártya kártyaterve



2. sz. ábra

TDM-gép

- a/ táprekesz
- b/ letapogató berendezés
- c/ II. rekesz
- d/ I. rekesz
- e/ kapcsolótábla
- f/ szelektor
- g/ az adatkereső berendezés regisztráló

