

UNITERM KATALÓGUS KÉSZÍTÉSE COMPUTERREL

Cseh László - Dezső László - Kutasy Kornél

Villamosenergiapari Kutatóintézet Számítóközpont

NIM Műszaki Dokumentációs és Fordító Iroda

Villamosenergiapari Kutatóintézet Számítóközpont

BEVEZETÉS

A könyvtári feldolgozási munkában computer alkalmazása korábban lehetetlennek látszott, mert a könyvtári munkában tulnyomó "szöveges" adat gépi feldolgozása az aránylag drága gépi idő mellett túlságosan költséges lett volna.

A computerok sebességének és memóriakapacitásának növekedése azonban bizonyos fordulatot jelent a szöveges információk feldolgozásával kapcsolatos probléma területén, mert a számítóközponton kívül, ún. szervező automatákon /lyukszalagos írógépeken/ előállított lyukszalagok információinak feldolgozása megfelelő gyorsműködésű nagy computer számára aránylag kis idejű igénybevételt jelent. A Razdan-3 szovjet gyártmányu computer belső memóriája két blokkban elrendezett $2 \times 16\ 384$ szó, egy szó 48 bit-ből áll. E nagy belső kapacitást max. 16 db dobmémória /7488 szó/dob/ és max. 16 db mágnesfilm egység /320 000 szó/egység/ egészíti ki és teszi gyakorlatilag korlátatlanná. A nagy működési sebesség /átlagosan 20 000 művelet/s/ mellett a gép másodpercenként 1000 karakter sebességgel tud szalagot beolvasni, és percenként 400-500 sor sebességgel adja ki sornyomatán az eredményt. Ugyanakkor természetesen szalagot is lyukaszthat 150 karakter/s sebességgel. Lehetőség van a számolási művelet, valamint a be- és kivitelt párhuzamos, tehát gyakorlatilag egy időben történő elvégzésére.

A gép fenti adatait figyelembe véve alakítottuk ki azt a tervet, hogy könyvtári anyag szak szerinti feltárásánál a géppel uniterm kártyákat /mondhatnánk táblákat/ állíttatunk elő, amikor a gép meglehetősen sok számadattal és aránylag kevés szöveges adattal fog dolgozni. Így a számolási, pontosabban sorbarendezi és átrendezi műveletek ideje olyan minimális lesz, hogy a rövid idő alatt is elvégezhető munka könyvtári szempontból ér annyit, amennyibe a gépi idő kerül.

Az alábbiakban a feladatot könyvtárosi, majd pedig számítástechnikai szemszögből szeretnénk ismertetni.

A FELADAT KÖNYVTÁROSI MEGFOGALMAZÁSA

Valamely gyűjtemény gyarapításakor a gyűjtemény soron következő dokumentumát bevezetjük egy leltárnaplóba, ahol megkapja a soron következő leltári számot, amely adott esetben /mechanikus raktározás esetében/ a helyszám szerepét is betöltheti. Ezt a számot a továbbiakban helyszámnak nevezzük.

A helyszámok megállapításával egyidejűleg sok esetben megoldható, hogy a feldolgozást végző könyvtáros a maga gyűjteményének külön szakozási táblázata alapján elvégezze a dokumentum szakozását és egyszerű jegyzékben rögzítse helyszámonként a helyszámhoz tartozó szakjelzeteket.

Feltételezzük, hogy a szakozási táblázat, amelyet a könyvtáros a saját gyűjteménye részére összeállított, megfelel az uniterm jellegű szakrendszernek, tehát egy olyan rendszernek, amelyben a tárgyszavak nem alárendeltek, hanem mellérendeltek /koordináltak/. Az ilyen egyenrangú tárgyszavakkal való szakozásnál természetesen nemcsak az előre megállapított tárgyszavakat lehet felhasználni, hanem a rendszer teljes jogú tagjává lehet felvenni újabb és újabb fogalmakat, mint pl. a gyártó cégek nevét vagy a termékek márkanevét, esetleg egyéb kategóriába tartozó fogalmakat.

Tekintettel arra, hogy kiindulási uniterm táblázatunk is tartalmazhat különféle kategóriákba tartozó tárgyszavakat /unitermeket/, vagyis tartalmazhat "tárgy"-szavakat, földrajzi fogalmakat, kronológiai fogalmakat stb., így a rendszerben már rögzített unitermek mellé az újonnan jelentkező vállalatneveket és egyéb unitermeket is felvehetjük, és uniterm kifejezés helyett a célszerűbb deskriptort használhatjuk. Tehát deskriptorainkat megszámozhatjuk és a szakozásnál ezt a deskriptorszámot - deskriptorkódot - használhatjuk szakjelzet gyanánt.

Figyelmünket irányítsuk most a két lényeges segédeszköze: szakrendi táblázatunkra, amelyet mindenkor kiegészítünk a jelentkező új fogalmakkal, valamint arra a jegyzékre, amelyben felsoroljuk a helyszámokat és a hozzájuk tartozó deskriptorkódokat.

Erre a két segédeszköze egyaránt jellemző, hogy a feldolgozás következő ütemére meg fognak változni. A szakrendi táblázat kiegészülhet új fogalmakkal, az a jegyzék pedig, amelyben felsoroljuk a helyszámokat és a hozzájuk tartozó deskriptorkódokat, más lesz, hiszen a feldolgozás következő üteme a helyszámoknak egy következő intervallumára, a beérkező dokumentumoknak egy következő csoportjára vonatkozik. Célunk az, hogy a mellérendelő természetű deskriptorok szerint csoportosítva megkapjuk a helyszámokat, mégpedig deskriptoroként olyan "töredékben", hogy az emelkedő sorrendben felsorolt odavágó helyszámok ne egy hasábian sorakozzanak egymás alatt, hanem tíz hasábian, úgy, hogy a hasáb sorszáma és a helyszám utolsó számjegye azonos legyen /l. ábra/. Az ismert uniterm elrendezéséről van szó, amely ezzel a felállítással megoldja, hogy különböző deskriptorok /ha

GOLDMAN ÖNJÁRÓTÁM

0	1	2	3	4	5	6	7	8	9
					1335				

1. ábra

Példa egy uniterm-kártyára

A deszkriptorhoz tartozó helyszámokat /jelen esetben 1335/ a gép az utolsó számjegyeknek megfelelő oszlopba írja

ugy tetszik, unitermek/ kártyáit egymás alá téve aránylag egyszerű összehasonlítással megtalálhassuk azokat a helyszámokat, amelyek az egymás alá rakott két, három vagy több uniterm kártyán "közösén" előfordulnak. Természetesen a kártyákra magát a deszkriptort akarjuk feliratozni, nem elégszünk meg a deszkriptorkóddal.

Ezzel tulajdonképpen a feladat "könyvtárosi" megfogalmazása meg is történt, mégis ide kívánczik egy apró fogás, amely a szakrendi táblázat vezetésénél, fejlesztésénél jó szolgálatot tesz. Rendszerint valamilyen meglévő szakrendi táblázatból szokás kiindulni, a példa kedvéért az ETO-ból válasszunk ki néhány bányászati osztályt, amelyekről azt állítjuk, hogy egymástól függetlenek, tehát mellérendeltek:

Vágatok, tárók biztosítása

Aknák biztosítása

Nagyméretű bányaterek biztosítása

Leművelt helyek biztosítása

Feltevésünk szerint ezeket megszámozzuk és így kapjuk a deszkriptorkódokat. Legyen deszkriptoraink számozása a következő:

35 Vágatok, tárók biztosítása

36 Aknák biztosítása

37 Nagyméretű bányaterek biztosítása

38 Leművelt helyek biztosítása

Ha most néhány vállalat vagy gyártmány márkanévét is deszkriptornak akarjuk használni, mint pl.

Országos Bányagépgyártó Vállalat
 Vörös Csillag Traktorgyár
 Ringsprung gyűri
 Goldman önjárótám,

akkor célszerű ezeket betűrendben ugrószámozással ellátni, hogy később jelentkező újabb márkanévek a betűrendbe besorolhatók legyenek, vagyis az ugrószámok között részükre is lehessen számokat kiutalni. Így pl. a következő számozást használhatjuk:

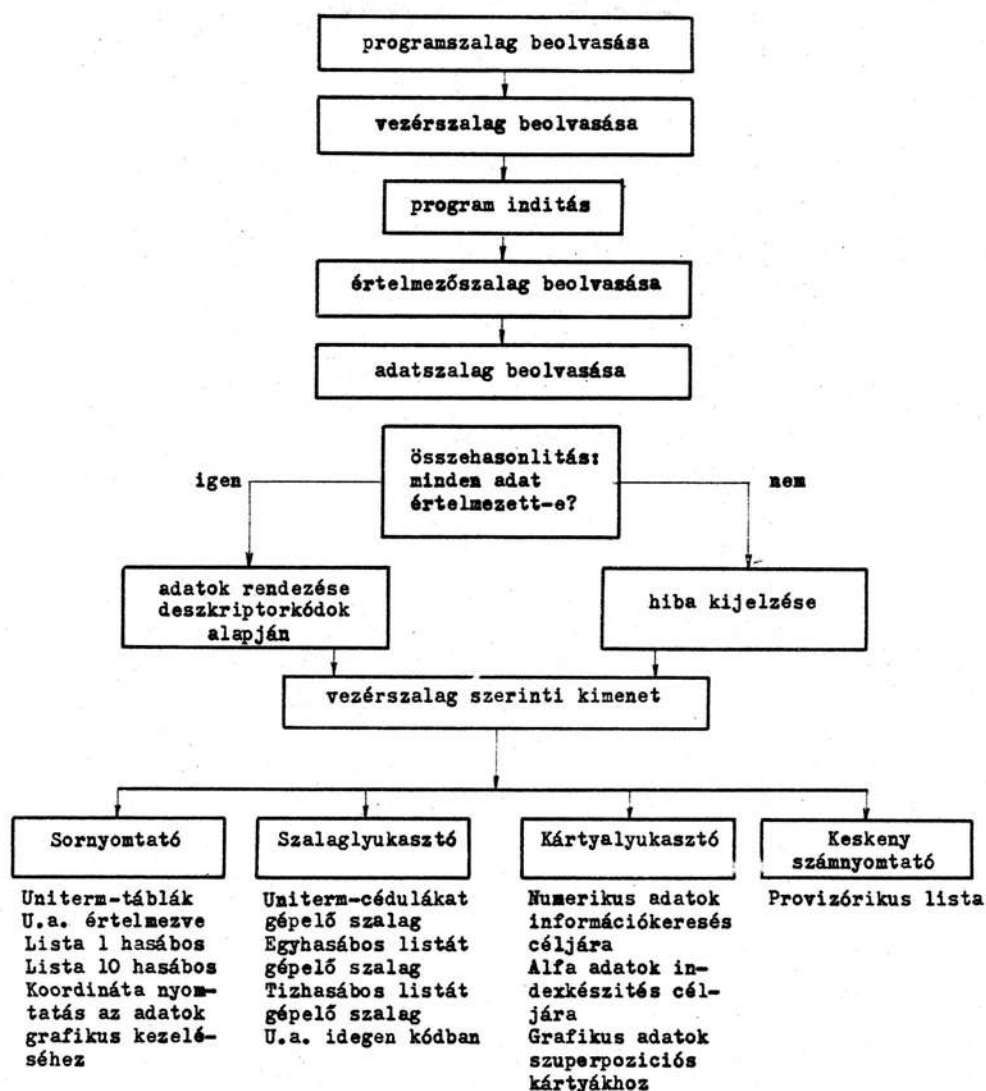
4250 Goldman önjárótám
 4255 Országos Bányagépgyártó Vállalat
 4260 Ringsprung gyűri
 4265 Vörös Csillag Traktorgyár

Az alábbiakban a könyvtáros már csak arról a nehézségről számolhatna be, hogy az uniterm-kártyán az adatok feltüntetése nem problémamentes. A feldolgozó munka elején ugyanis felfektetünk megfelelő számú uniterm-kártyát, elvégezzük a szakozást azáltal, hogy az egymás után kézbevevett dokumentumok helyszámát a vonatkozó kártya megfelelő oszlopába beírjuk. Ilyenkor indokolatlannak látszik a deszkriptor mellett még egy deszkriptorkód szerepeltetése is, hiszen magát a kártyát használjuk. Sajnos a feldolgozó rendszer alakítása, kiegészítése alkalmával minduntalan előfordul, hogy már felfektetett kártyákra korábbi helyszámot akarunk majd felírni, ami sérti az uniterm-kártya rendjét, megzavarja hasábonként a helyszámok emelkedő sorrendjét. Ezen a nehézségen is segít a computeres feldolgozás, amelynek lényege, hogy az először elkészített uniterm táblákat a következő alkalommal elkészült táblák teljes mértékben feleslegessé tesszik. Tehát a gép mindenkor új táblákat nyomtat, amelyeken a később felbukkanó helyszámok a saját helyükre kerülnek.

A FELADAT SZÁMITÁSTECHNIKAI MEGFOGALMAZÁSA

Számitástechnikai szempontból a munka célszerű elnevezése: adatrendezés. Az adatrendezési munkának az a célja, hogy egy vagy több adatszalagról beolvasott számcsoportokból számpárokat állítson elő és a számpárokat valamelyik számtípus szerint rendezze, a vezérszámokat az értelmezőszalagról beolvasott értelmezéssel egészítse vagy cserélje ki, végül kimenetként a sornyomatón adjon valamilyen jegyzéket vagy táblát /pl. uniterm táblát/, vagy szalaglyukasztó segítségével olyan szalagot állítson elő, amely lyukszalagos írógépen olvasható /kiiratható/, vagy lyukkártyát lyukasszon valamilyen elgondolás szerint. A program blokkismája a 2. ábrán látható.

A Razdan-3 computer rendkívül nagy operatív memóriája lehetővé teszi, hogy kb. 3000 helyszámot - helyszámonként kb. 6 deszkriptorkódot - tartalmazó adatszalagot külső tároló /mágnesfilm, mágnesle-



2. ábra
Programvázlat

mez/ felhasználása nélkül tudunk feldolgozni, ami a feldolgozási időt nagymértékben csökkenti.

A következőkben a gépi feldolgozáshoz szükséges lyukszalagokat ismertetjük részletesen.

A szalagok lyukasztása az IBM 8 csatornás kódnak megfelelően történik.

Két számtípus van, um. helyszám és deszkriptorkód.

Az adatszalagra lyukasztott helyszám azzal jellemezhető, hogy "kocsi-vissza" /CR/ után és "szóköz" /SP/ előtt áll, mégpedig vagy magában mint legfeljebb 5-jegyű szám, vagy alsó és felső határt jelölve, amikor is e két határszámot mínusz-jel köti össze. A helyszám természetére nézve sorszám, adott intervallumban a pozitív egész számok - többnyire - hiánytalan sora.

Az adatszalagra lyukasztott deszkriptorkód azzal jellemezhető, hogy előtte SP, utána vagy SP vagy CR áll. A deszkriptorkód rendszerint ugrószám. Tartalmazhat tizedesvesszőt. Sorrendjét az értelmezőszalagon rögzített sorrend szabja meg. Helyértékigénye hat, vagyis öt szignifikáns számjegy és tizedesvessző /3. ábra/.

37 4250 4265

Nagyméretű bányaterek biztosítása
Goldman-önjárótám segítségével.
Bp. Vörös Csillag Traktorgyár,
1967. 62 p. 1335

3. ábra

Példa egy olyan címleírásra, amely előtt deszkriptorkódok vannak, utána helyszám /itt 1335/ következik. Az ilyen szövegeket tartalmazó szalag jellege szerint értelmező szalag, pontosabban szöveg-szalag

Az adatszalag végén stop-kód /STP/ van.

Az értelmező szalagon, jellegétől függően, valamelyik típusu szám áll és a hozzátartozó szöveges értelmezés. Kétféle értelmező szalag lehet, um. szöveg-szalag és deszkriptor-szalag /4. ábra/.

36 Aknák biztosítása

37 Nagyméretű bányaterek biztosítása

4. ábra

Példa egy olyan értelmező szalagra,
amely a deszkriptorkódokhoz tartozó deszkriptorokat sorolja fel
/deszkriptor-szalag/

A szöveg-szalag tartalmaz bibliográfiai jellegű szöveget, amelynek végén tabulátor /TAB/ kód után és CR előtt található a helyszám /5. ábra/.

1330 12 33 35 75 4211

1331 2 4250

1332 23 24 27 4100

1334 38 4260 4250

1335 37 4250 4265

5. ábra

Példa az adatszalagon szereplő helyszámok
és szakjelzetek kombinációjára.

A computer adatrendező munkája mindenkor ilyen numerikus adatokkal
megy végbe

A deszkriptor szalagon CR és TAB után áll a deszkriptorkód, CR után annak értelmezése. Ez egy vagy több sorra terjedhet.

Mindkét típusú értelmező szalag végén STP kód van.

Az adatrendezés szervezésében az alábbi variációkat kell figyelembe venni:

értelmező szalag:/tehát vagy szöveg-szalag a 3. ábra szerint,
vagy deszkriptor-szalag a 4. ábra szerint/, esetleg nincs is, vagy
csak deszkriptor szalag van /pl. uniterm katalógus készítésekor/,
vagy van szöveg szalag is /pl. nyomtatott katalógus készítésekor/;

adatszalag:/5. ábra/ a rendezési cél lehet deszkriptorkód szerinti csoportosítása a helyszámoknak /ez a tulnyomó/, vagy a deszkriptorkódok helyszámok szerinti csoportosítása;

rendezés: ld. az előző pontot;

értelmezés: módja több tényezőtől függ aszerint, hogy milyen a rendezés célja és milyen kimenetre van szükség. Így az értelmezés történhet a sornyomtató részére a sornyomtató kódjának figyelembevételével, vagy szalaglyukasztás részére az eredetivel azonos kódban, esetleg más kódban, más output készülék részére. Az értelmezés kiter-

jedhet az egyik vagy a másik fajta adatokra, vagy mindkét fajta adatokra;

Kimenet: több variáció lehetséges egyrészt kimenő kód szerint, másrészt kimenő készülék szerint, harmadrészt a kimenet rendeltetése, típusa szerint. Így a kód lehet a sornyomatató kódja, a szalaglyukasztó kódja, a kártyalyukasztó kódja. A készülék lehet sornyomatató, szalaglyukasztó vagy kártyalyukasztó stb. A kimenet típusa lehet uniterm vagy szuperpozíciós tábla, vagy egyszerű felsorolás, esetenként 10 hasámban. Szalaglyukasztáskor lehet szintén uniterm vagy felsorolás, esetleg több hasámban, végül kártyalyukasztó esetben lehet valamilyen alfanumerikus kártyaterv szerinti, de lehet geometrikus /koordináta hálózathoz megfelelő/ kimenet is.

A számos variációra számítani kell a program szervezésénél.

ÖSSZEFOGLALÁS

Az adatrendezés elnevezésű munka számítástechnikai szempontból így jellemezhető: a program alkalmas arra, hogy a gépbe beolvassunk egy értelmező szalagot, amely a deszkriptorkódokat és a deszkriptorokat sorolja fel; alkalmas arra is, hogy egy másik értelmező szalagot olvassunk be a gépbe, amely helyszámokat és a helyszámokhoz tartozó adatokat /szöveges adatokat/ tartalmazza, pl. címleírásokat, amelyeket adott esetben akár referátumok is követhetnek. Alkalmas továbbá a program arra, hogy beolvasson adatszalagot, amely helyszámok és deszkriptorkódok összetartozását örökíti meg. Miután a program ellenőrzi, hogy a beolvasott helyszámok egy konkrétan megadott helyszám-intervallumon belül vannak-e, továbbá ellenőrzi, hogy a beolvasott deszkriptorkódok mindegyike szerepel-e a már beolvasott értelmező szalag deszkriptorkódjai között, előállít a helyszámokból és a deszkriptorkódokból számpárokat, ezeket rendezi, és a deszkriptorkódokként sorbarendezett helyszámok birtokában a következő műveletek valamelyikét végezheti el a vezérszalagtól függően:

A deszkriptorkódokat kicseréli szövegre /szám helyett szöveget ad/;

Deszkriptoronként a helyszámokat

sornyomatón kinyomtatja,

sornyomatón uniterm-kártyaként kinyomtatja,

szalagra lyukasztja azzal a céllal, hogy a szalagról másutt listát vagy

uniterm-táblát lehessen gépelni;

A helyszámkódokat kiegészíti a hozzájuk tartozó szöveggel /ha volt un. szöveg-szalag/;

Desszkriptoronként ezeket a szövegeket

sornyomtatón kinyomtatja,

szalagot lyukaszt azzal a céllal, hogy a szalagról másutt /ahol olcsóbb a gépidő/ listát, vagy

előre megadott sorszáma tördelt sokszorosítási tükröt. gépeljenek.

A sokoldaluan használható program még alkalmassá tehető arra is, hogy a bibliográfiai adatokat tartalmazó értelmező szalagról a neveket és a nevek kezdőbetűit felismerje, leválassza és a sokszorosítási tükröz "szerkesztésével" kialakuló lapszámozással kiegészítve névmutatót készítsen egy-egy kiadványhoz.

A program nélküli kiadványok lyukszalagos írógépekkel való előállítás nem automatizálható, az egyszerű gépelés automatikus átmásolása és ilyen formában való további kezelése meglehetősen sok időt követel a lyukszalagos írógép kezelőjétől. Ugyanakkor a gondosan előállított lyukszalag - megfelelő korrektúrázás után - computerbe olvasható a fentiek szerint és néhány perces gépidő árán olyan szalag nyerhető, amely a kész sokszorosítási tükröt adja sornyomtatón vagy olvasásra kész lyukszalagon. Utóbbi esetben a lyukszalagos írógépek gazdagabb jelkészlete /nagybetűk és kisbetűk alkalmazásának lehetősége/ jelent előnyt.



CSEH, L. - DEZSÓ, L. - KUTASY, K.: Producing Uniterm catalogue by computer

A program has been prepared for a Soviet-made Razdan-3 digital computer which is capable of processing library catalogue entries. The program is many-sided. One of its possibilities is that the computer ignores the entries themselves and knows only the "classification scheme /tables/" and the groups of the related call numbers and classification notations. In this case the computer sorts the numerical data according to the classification notations belonging to the call number, then it will draw forth the descriptor /i.e. the textual equivalent of the notation/, from its store and will produce a Uniterm table so that it indicates not only the mere notation per item /by which the sorting and ordering operation was performed/ but also the textual equivalent of the notation /descriptor/.

The other possibility: Uniterm tables are not produced by the printer unit, but the output is a punched paper tape applicable to producing Uniterm cards on typewriter. In this case, for example an 8-channel paper tape may produce both upper and lower case letters, diacritical marks and symbols.

Further possibilities are provided by punched cards as output, and incidentally by such geometric formations which permit the production of Peek-a-boo cards.

Besides the "classification scheme", catalogue entries may also be fed into the machine. In this case, after ordering the above mentioned numerical data pairs /i.e. call number and notation/, the machine is able to substitute not only descriptors for notations, but also entries for call numbers. Thus the ordering of data results in the sequence order of a "printed catalogue". Subsequently - now that it does not produce Uniterm cards - substituting headings for notations, and entries for call numbers, the computer may produce - as output - a paper tape which is applicable for the machine setting of a publication.



ЧЕХ, Л. - ДЭЖЭ, Л. - КУТАШИ, К.: Составление каталога УниTERM
с помощью ЭЦВМ

Для электронной вычислительной машины Раздан-3 советского производства создана такая программа, которая пригодна для составления библиографических описаний для библиотек. Программа многосторонняя. Одна из возможностей заключается в том, что библиографическими описаниями также и не занимаются, а только приводятся "таблица специальностей" и собранные воедино группы номеров мест и специальностей. В этом случае цифровые данные распределяются согласно шифрам специальностей, затем из памяти отбирается текстовое соответствие шифру специальности (дескриптор) и составляется таблица УниTERM таким образом, что в ее шапке фигурируют не пустые шифры. (С которыми производилось распределение данных), а текст, относящийся к шифру специальности (дескриптор).

Вторая возможность состоит в том, что такие таблицы УниTERM печатает не строко-печатающее устройство машины, а на выходе вылается перфолента, которая пригодна для заполнения бланков УниTERM. В этом случае (напр. с 8-канальной перфоленты) можно изобразить строчные и заглавные буквы, орфографические знаки, а также знаки препинания.

Новые возможности истекают из того, что выходная перфокарта представляет собой такую геометрическую формацию, с которой можно изготовить световую перфокарту.

Кроме "таблицы специальностей", в машину можно ввести и библиографические описания. В этом случае после вышеупомянутого распределения пар цифровых данных машина может заменить на текстовую формацию не только шифры специальностей, но и номера мест (т.е. на библиографические описания), и таким образом распределение данных создает порядок построения "печатного каталога". Затем заменив шифры специальностей на названия граф и номера мест на библиографические описания, ЭЦВМ (теперь она изготавливает не карты УниTERM) может перформировать выходную перфоленту, пригодную для набора всего издания.

--X--

CSEH, L. - DEZSÓ, L. - KUTASY, K.: Herstellung von Uniterm-Katalogen mit Computer

Es wurde ein Programm für den sowjetischen Razdan-3 Digital-Computer zusammengestellt, welches für eine vielseitige Bearbeitung von Bibliotheks-Titelaufnahmen geeignet ist. Eine Möglichkeit besteht darin, dass sich der Computer mit Titelaufnahmen gar nicht beschäftigt, er kennt nur die "Sachkatalogisierungs-Tabelle", sowie die zusammengehörenden Signatur - DK-Nummer Gruppen. In diesem Fall ordnet er die numerischen Daten nach den zu den Signaturen gehörenden DK-Nummern: er nimmt dann aus seinem Speicher der DK-Nummern äquivalenten Deskriptoren hervor und stellt eine Uniterm-Tabelle in der Weise zusammen, dass in dieser als Kopf nicht die DK-Nummer allein /mit Hilfe deren die Datenordnung durchgeführt wurde/, sondern auch der zur DK-Nummer gehörende Text - Deskriptor - dargestellt wird.

Eine andere Möglichkeit besteht darin, dass die Uniterm-Tabellen nicht vom Reihendrucker des Maschinenparks gedruckt werden, sondern der Ausgang gibt einen Lochstreifen, welcher für das Maschinenschreiben von Uniterm-Karten geeignet ist. In diesem Fall /z.B. auf einem Lochstreifen mit 8 Kanälen/ können auch grosse und kleine Buchstaben, Buchstaben mit spezieller Punctuation und auch Symbole hergestellt werden.

Weitere Möglichkeiten entstehen dadurch, dass die Ausgangs-Lochkarte gegebenenfalls eine solche geometrische Formation bildet, von welcher eine Sichtlochkarte verfertigt werden kann.

Ausser der "Sachkatalogisierungs-Tabelle" können auch die Titelaufnahmen in die Maschine eingetragen werden. In diesem Fall kann die Maschine nach der obenerwähnten Ordnung der numerischen Datenpaaren nicht nur die DK-Nummern mit Text-Formationen austauschen, sondern auch die Signaturen /nämlich mit den Titelaufnahmen/ und so stellt die Datenordnung das Aufbausystem eines "gedruckten Katalogs" her. Danach kann der Computer - da er nunmehr keine Uniterm-Karten herstellt - nach Ersetzung der DK-Nummern durch Köpfe und die der Signaturen durch Titelaufnahmen einen für den Druck einer ganzen Publikation geeigneten Lochstreifen als Ausgang löchern.