*Alkalmazott Matematikai Lapok* **9** (1983) 13—21

# A FÜGGŐSÉGEK ÉS AZ INDIVIDUUMOK SZÁMA KÖZÖTTI KAPCSOLAT ÖSSZETETT ADATRENDSZEREK ESETÉN

DEMETROVICS JÁNOS FÜREDI ZOLTÁN KATONA GYULA

Budapest

Egy korábbi dolgozatunkban azt a feladatot vizsgáltuk, hogy egy adott kulcsrendszerrel rendelkező adatbázisban mennyi az individuumok minimális száma. A jelen dolgozatban részben ezzel a problémával, részben egy természetes analogonnal foglalkozunk. Azt is vizsgálni fogjuk ugyanis, hogy mennyi az individuumok minimális száma, ha az adatbázisban a függőségek rendszere is adott.

Sokszor hasznos az, ha az összetett adatbázis bizonyos paramétereit ki tudjuk számítani, vagy meg tudjuk becsülni a rész-adatbázisok megfelelő paramétereiből. Ezt kíséreljük meg elvégezni az említett paraméterre, az individuumok minimális számára. Erre jelenleg csak akkor vagyunk képesek, ha a nagy adatbázis egymástól független kisebbekből van felépítve.

## 1. Bevezetés

Az adatbázis struktúrájának a felhasználó számára legfontosabb jellemzője az, hogy a tárolt adatok közötti összefüggések milyen formában vannak jelen az adatbázisban. Mivel egy adatbázis struktúráját a rekordtípusok és a rekordtípusok közötti kapcsolatok alkotják, ezért az adatmodelleket általában aszerint osztályozzuk, hogy a rekordtípusok között milyen kapcsolat definiálása megengedett. Ennek alapján lényegében két adatmodellt különböztetünk meg:

— az egyik az ún. *hálózatos adatmodell,*
— a másik az ún. *relációs adatmodell.*

Rekordtípus vagy más néven adatbázis megadása matematikai szempontból nem más, mint egy „értelmezési tartomány" és egy „értékkészlet" definiálása. Az így kapott rekordtípus előfordulásai az adott értelmezési tartományon értelmezett, értékeit az adott értékkészletből felvevő függvények. Az értelmezési tartomány elemeit attributumoknak nevezik. Egy adatbázis egy rekordtípusának egy adott pillanatban létező előfordulásai tehát egy mátrixot alkotnak. Ennek sorai az egyes előfordulások, oszlopai pedig az egyes attributumokon az előfordulások értékei.

Egy korábbi dolgozatunkban [2] vizsgáltuk azt a feladatot, hogy egy adott kulcsrendszerrel rendelkező adatbázisban mennyi az individuumok minimális száma. A jelen dolgozatban részben ezzel a problémával, részben egy természetes analogonnal foglalkozunk. Azt is vizsgálni fogjuk ugyanis, hogy mennyi az individuumok minimális száma, ha az adatbázisban a függőségek rendszere is adott.

Egy adatbázis sokszor természetes módon épül fel több adatbázisból. Máskor egy adatbázist mesterséges módon igyekszünk felbontani kisebb adatbázisokra. Mindkét esetben hasznos, ha az összetett adatbázis bizonyos paramétereit ki tudjuk számítani, vagy meg tudjuk becsülni a rész-adatbázisok megfelelő paramétereiből. Ezt kíséreljük meg elvégezni az említett paraméterre, az individuumok minimális számára. A kutatás jelenlegi állásában erre is csak akkor vagyunk képesek, ha a nagy adatbázis egymástól független kisebbekből van felépítve. Természetesen ugyanerre szükség volna azon esetben is, amikor az összeépítés bonyolultabb.

A 4. fejezetben a fentiekkel kapcsolatos további gondolatok vannak.

## 2. Definíciók, alapfogalmak

Az adatbázis egyik legfontosabb fogalma az E. F. CODD [3] által definiált funkcionális függés. Az előzőekből megállapíthatjuk, hogy egy relációs adatbázis pillanatnyi tartalmát mátrixnak tekinthetjük, ezért a funkcionális függést mátrixokra definiáljuk.

Lássuk most a pontos matematikai definíciókat. Egy *adatbázist* úgy tekintünk, mint egy $m$ sorból és $n$ oszlopból álló mátrixot, amelynek értékei nemnegatív egész számok, és amelynek minden sora különböző. A sorok felelnek meg az individuumoknak, az oszlopok az attributumoknak. Az oszlopok egy halmaza *kulcs*, ha nincs két olyan különböző sor, amelyek megegyeznek ezekben az oszlopokban. Vagyis a kulcs oszlopaiban álló értékek egyértelműen meghatározzák a sort. Ennél általánosabban, ha $A$ és $B$ az oszlopok egy-egy halmaza, és minden két olyan sor, amely megegyezik $A$-ban, az megegyezik $B$-ben is, akkor azt mondjuk, hogy $A$-ból *következik* $B$ (jelölésben $A \to B$). Az ilyen $(A, B)$ párokat függőségeknek nevezzük. Jól látható, hogy ha $A$ egy kulcs és $X$ az összes oszlopok halmaza, akkor $A \to X$. Egy $A$ kulcsot *minimális kulcs*nak nevezünk, ha nincs olyan $B$ kulcs, melyre $B \subset A$, $B \neq A$ teljesülne. A minimális kulcsok rendszerét $\mathcal{K}$-val jelöljük. Ismeretes [4], hogy $\mathcal{K}$ akkor és csak akkor lehet valamely adatbázis kulcsainak rendszere, ha nem üres *Sperner-rendszer,* azaz nincs két különböző $A, B$ eleme, amelyekre $A \subset B$ fennállna.

Legyen az $M$ adatbázishoz hozzárendelve a következő $X$ részhalmazából $Y$ részhalmazába leképező függvény:

$$\mathcal{L}_M(A) = \{b : A \to \{b\}.\} \tag{2.1}$$

Könnyen látható, hogy ez az $\mathcal{L}_M$ függvény egyrészt egyértelműen meghatározza a függőségek teljes rendszerét $(A \to B \Leftrightarrow B \subseteq \mathcal{L}_M(A))$ másrészt rendelkezik a következő tulajdonságokkal:

$$A \subseteq \mathcal{L}(A) \tag{2.2}$$

$$A \subseteq B \Rightarrow \mathcal{L}(A) \subseteq \mathcal{L}(B), \tag{2.3}$$

$$\mathcal{L}(\mathcal{L}(A)) = \mathcal{L}(A), \tag{2.3a}$$

a fenti három tulajdonsággal rendelkező $\mathcal{L}$ függvényeket *lezárásoknak* nevezzük. Ismeretes, hogy minden $\mathcal{L}$ lezáráshoz van olyan adatbázis, amelyben a függőségek által generált $\mathcal{L}_M$ függvény éppen $\mathcal{L}$-lel egyenlő.

Legyen $\mathcal{K}$ az $n$-elemű halmaz részhalmazainak egy *Sperner-rendszere.* Mint fent említettük, létezik olyan adatbázis, amelyben ez a minimális kulcsok rendszere. Legyen ezen adatbázisok közül a minimális sorból állónak $s(\mathcal{K})$ sora. [1]-ben találhatók $s(\mathcal{K})$-ra vonatkozó eredmények néhány egyszerű $\mathcal{K}$ rendszer esetén. A fentiekkel analóg módon definiálhatjuk $s(\mathcal{L})$-et: Egy adott $\mathcal{L}$ lezáráshoz található adatbázisok közül a minimális sorúak sorszámaként.

Ha $\mathscr{L}_1$ egy $n_1$ elemű $X_1$ halmazon egy lezárás, $\mathscr{L}_2$ pedig egy $n_2$ elemű $X_2$-n $(X_1 \cap X_2 = \emptyset)$, akkor definiáljuk az $n_1+n_2$ elemű $X_1 \cup X_2$-n az $\mathscr{L}_1$ és $\mathscr{L}_2$ lezárások $\mathscr{L}_1 \times \mathscr{L}_2$ direkt szorzatát a következőképpen:

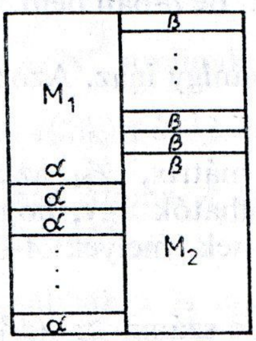
$$(2.4) \qquad (\mathscr{L}_1 \times \mathscr{L}_2)(A) = \mathscr{L}_1(A \cap X_1) \cup \mathscr{L}_2(A \cap X_2).$$

Hasonlóan, ha $X_1$-en, illetve $X_2$-n csak a minimális kulcsok $\mathscr{K}_1$, illetve $\mathscr{K}_2$ rendszere adott, akkor $\mathscr{K}_1 \times \mathscr{K}_2$ jelenti $X_1 \cup X_2$ azon részhalmazainak rendszerét, amelyek egy $\mathscr{K}_1$-beli és egy $\mathscr{K}_2$-beli uniójaként állnak elő.

Ezek után kimondhatjuk főeredményünket:

2.1. Tétel.
$$s(\mathscr{L}_1 \times \mathscr{L}_2) = s(\mathscr{L}_1) + s(\mathscr{L}_2) - 1.$$

*Bizonyítás.* 1. Bizonyítsuk először a $\leqq$ egyenlőtlenséget. Legyen $\mathbf{M}_1$ egy olyan $s(\mathscr{L}_1)$-szer $n_1$-es mátrix, amelyre $\mathscr{L}_{M_1} = \mathscr{L}_1$, azaz amely $\mathscr{L}_1$-et a legkevesebb sorral realizálja. Hasonlóan, legyen $\mathbf{M}_2$ az $\mathscr{L}_2$ lezárást legkevesebb sorral realizáló mátrix. Ezek segítségével fogunk egy $s(\mathscr{L}_1)+s(\mathscr{L}_2)-1$ sorú és $n_1+n_2$ oszlopú $\mathbf{M}$ mátrixot konstruálni, amelyre $\mathscr{L}_M = \mathscr{L}_1 \times \mathscr{L}_2$. $\mathbf{M}$ konstrukcióját a következő séma szemlélteti:




*1. ábra*

Azaz $\mathbf{M}_1$-et és $\mathbf{M}_2$-t úgy helyezzük egymás mellé, hogy csak egy sorban találkozzanak, $\mathbf{M}_1$ utolsó sora, $\alpha$ legyen $\mathbf{M}_2$ első sorával, $\beta$-val egy vonalban. Ezután $\mathbf{M}_2$ többi sora mellé $\alpha$-t, $\mathbf{M}_1$ többi sora mellé $\beta$-t írjunk. Triviális, hogy $\mathbf{M}$-nek $s(\mathscr{L}_1)+s(\mathscr{L}_2)-1$ sora és $n_1+n_2$ oszlopa van, tehát csak azt kell bizonyítani, hogy az $\mathbf{M}$-hez tartozó $\mathscr{L}_M$ lezárás $\mathscr{L}_1 \times \mathscr{L}_2$. Vagyis, ha $A \subset X = X_1 \cup X_2$, $a \in X$, akkor

$$(2.5) \qquad a \in \mathscr{L}_M(A) \Leftrightarrow a \in \mathscr{L}_1(A \cap X_1) \cup \mathscr{L}_2(A \cap X_2).$$

A szimmetria miatt feltehetjük, hogy $a \in X_1$. Ekkor (25) úgy is fogalmazható, hogy a

(2.6a) $a \in \mathscr{L}_1(A \cap X_1) \Rightarrow \mathbf{M}$ minden olyan két sora, amely $A$-ban megegyezik, megegyezik $a$-ban is

és

(2.6b) $a \notin \mathscr{L}_1(A \cap X_1) \Rightarrow \mathbf{M}$-nek van két olyan sora, melyek $A$-ban megegyeznek, de $a$-ban nem

feltételeknek együtt kell teljesülnie. Bizonyítsuk be tehát ezeket. Ha $a \in \mathscr{L}_1(A \cap X_1)$, ekkor $\mathbf{M}_1$ definíciója szerint $\mathbf{M}_1$ minden olyan két sora, amely $A \cap X_1$-en megegyezik, $a$-n is megegyezik. Ha $\mathbf{M}$ kiválasztott két sora (ami $A$-ban megegyezik) mindkettő $\alpha$-val kezdődik, nincs mit bizonyítanunk. Ha egyik se kezdődik $\alpha$-val, akkor $\mathbf{M}_1$ fenti tulajdonságából következik az $a$-ban való megegyezés. Ha egyik $\alpha$-val kezdődik, a másik nem, akkor ismét $\mathbf{M}_1$ fenti tulajdonsága elegendő, hiszen $\alpha$ az $\mathbf{M}_1$ egy sora.

Térjünk most rá (2.6b) bizonyítására. Tegyük fel, hogy $a\notin\mathcal{L}_1(A\cap X_1)$. $\mathbf{M}_1$-nek definíció szerint létezik két olyan sora, melyek $A\cap X_1$-ben egyeznek, de $a$-ban nem. Az ezeknek megfelelő két sor $\mathbf{M}$-ben kielégíti (2.6b) jobb oldalát.

2. A $\geqq$ egyenlőtlenség bizonyításához vegyünk egy tetszőleges $\mathbf{M}$ mátrixot, melyre $\mathcal{L}_{\mathbf{M}}=\mathcal{L}_1\times\mathcal{L}_2$ és bizonyítsuk be, hogy $\mathbf{M}$ sorainak száma $\geqq s(\mathcal{L}_1)+s(\mathcal{L}_2)-1$. Ennek bizonyításához két lemmára lesz szükségünk. Az $\mathbf{M}$ mátrix $\mathcal{L}_1$-nek megfelelő első $n_1$ oszlopát jelöljük $\mathbf{M}_1$-gyel, a maradékot $\mathbf{M}_2$-vel.

2.1. Lemma. $\mathcal{L}_{\mathbf{M}_2}=\mathcal{L}_2$.

*Bizonyítás.* Legyen $A\subset X_2$, $a\in X_2$. Lássuk be először, hogy ha $a\in\mathcal{L}_2(A)$, akkor $a\in\mathcal{L}_{\mathbf{M}_2}(A)$. A feltételből és $\mathbf{M}$ definíciójából következik, hogy ha $\mathbf{M}$ két sora megegyezik $A$-ban, akkor $a$-ban is. Ez természetesen az $\mathbf{M}_2$ részmátrixban is igaz. Megfordítva, ha $a\notin\mathcal{L}_2(A)$, akkor $a\notin(\mathcal{L}_1\times\mathcal{L}_2)(A)$, így $\mathbf{M}$-nek van két olyan sora, amelyek $A$-ban megegyeznek, de $a$-ban nem. Ugyanez a két sor $\mathbf{M}_2$-ben is megfelel, azaz $a\notin\mathcal{L}_{\mathbf{M}_2}(A)$.

Természetesen $\mathcal{L}_{\mathbf{M}_2}=\mathcal{L}_1$ ugyanígy igaz. Azonban $\mathbf{M}_1$-re egy kissé erősebb állítást fogunk használni:

2.2. Lemma. Legyen $\mathbf{N}$ egy mátrix, $\mathcal{L}_{\mathbf{N}}$ az általa generált lezárás. Tegyük fel, hogy $\mathbf{N}$ sorai $k$ osztályba sorolhatók úgy, hogy ha $a\notin\mathcal{L}_{\mathbf{N}}(A)$, akkor található 2, egy osztályba tartozó sora $\mathbf{N}$-nek, melyek $A$-ban megegyeznek, de $a$-ban nem. Ekkor

$$\mathbf{N}\text{ sorainak száma }\geqq s(\mathcal{L}_{\mathbf{N}})+k-1. \tag{2.7}$$

*Bizonyítás.* A bizonyítást $k$-ra vonatkozó indukcióval végezzük. Ha $k=1$, (2.7) definíció szerint igaz. Tegyük fel tehát, hogy $k$-ra igaz és bizonyítsuk be $k+1$-re. $\mathbf{N}$ sorai legyenek $k+1$, a lemma feltételeinek megfelelő osztályba sorolva. Ha $\mathbf{N}$-nek konstans oszlopai lennének, hagyjuk el ezeket. Ez módósitott mátrix, amelyet továbbra is $\mathbf{N}$-nel jelölünk, a lemma feltételeit változatlaul teljesíti, $\mathcal{L}_{\mathbf{N}}(\emptyset)=\emptyset$. Változtassuk meg $\mathbf{N}$ értékeit úgy, hogy két különböző osztályban egyáltalán ne legyen azonos érték, viszont egy osztályon belül az azonosság-különbözőség ne változzon (pl. az $i$-edik osztályban levő sor egy $j$ eleme helyett írjunk $j\cdot p^i$-t, ahol $\cdot p$ egy olyan prímszám, ami minden $\mathbf{N}$-ben szereplő értéknél nagyobb). Az így keletkezett mátrixot jelöljük $\mathbf{N}'$-vel. Gondoljuk át, hogy

$$\mathcal{L}_{\mathbf{N}'}=\mathcal{L}_{\mathbf{N}}. \tag{2.8}$$

Tegyük fel, hogy $a\in\mathcal{L}_{\mathbf{N}}(A)$, azaz, ha $\mathbf{N}$ két sora $A$-ban azonos, akkor $a$-ban is. Ekkor, ha $\mathbf{N}'$ két sora azonos $A$-ban (akkor $A\neq\emptyset$ esetén $\mathbf{N}$ azonos osztályából valók), akkor ugyancsak egyformák lesznek $a$-ban is. Így $a\in\mathcal{L}_{\mathbf{N}'}(A)$. Másrészt, ha $a\notin\mathcal{L}_{\mathbf{N}}(A)$, akkor $\mathbf{N}$-nek van két, egy osztályba eső sora, amelyek $A$-ban egyeznek, $a$-ban nem. Mivel egy osztályban vannak, ugyanez $\mathbf{N}'$-ben is teljesülni fog rájuk. Következésképpen $a\notin\mathcal{L}_{\mathbf{N}'}(A)$.

Legyen $\mathbf{N}'$ első és második osztályának egy-egy sora $\gamma=(\gamma_1, \ldots, \gamma_u)$, illetve $\delta=(\delta_1, \ldots, \delta_u)$. Hagyjuk el a mátrixból $\gamma$-t és írjunk az első osztály $i$-edik oszlopában szereplő minden $\gamma_i$ helyett $\delta_i$-t. Az így kapott $\mathbf{N}''$ mátrixra igazoljuk most az

$$\mathscr{L}_{\mathbf{N}''} = \mathscr{L}_{\mathbf{N}'} \tag{2.9}$$

állítást. Tegyük fel, hogy $a\in\mathscr{L}_{\mathbf{N}'}(A)$ és legyen $\mathbf{N}''$ két sora $A$-ban azonos. Ha ez a két sor $\mathbf{N}'$ szerinti azonos osztályokban van, akkor a két sor $a$-ban való azonossága következik. Ha nem, akkor csak a volt első és második osztály egy-egy sora lehet, $\mu''$, illetve $\gamma''$. Itt a két sor $i$-edik eleme csak akkor lehet azonos, ha mindkettő $\delta_i$. De akkor a megfelelő $\mathbf{N}'$-beli sorban, $\mu'$-ben $\gamma_i$ áll az $i$-edik helyen. Azaz, $\mu'$ az $A$-ba eső részében megegyezik $\gamma$-val. $a\in\mathscr{L}_{\mathbf{N}'}(A)$-ból következik, hogy e sor $a$-adik eleme $\mathbf{N}'$-ben $\gamma_a$, $\mathbf{N}''$-ben $\delta_a$. Hasonlóan látható, hogy a $\nu'$ sor $a$-adik eleme $\delta_a$, vagyi $\mu''$ és $\nu''$ megegyezik az $a$-adik helyen. $a\in\mathscr{L}_{\mathbf{N}''}(A)$.

Megfordítva, tegyük fel, hogy $a\notin \mathscr{L}_{\mathbf{N}'}(A)$, azaz $N'$-nek létezik két sora, amelyek $A$-ban egyeznek, $a$-ban nem. Ezek természetesen azonos osztályba fognak esni. Ez a két sor változatlanul megfelel $\mathbf{N}''$-ben is. $a\notin \mathscr{L}_{\mathbf{N}''}(A)$ következik. Ebből (2.9)-en kívül az is látható, hogy $\mathbf{N}''$-re a lemma feltevése már $k$-val teljesül, alkalmazható az idukciós feltevés:

$$(\mathbf{N} \text{ sorainak száma})-1 = \mathbf{N}'' \text{ sorainak száma } \geqq s(\mathscr{L}_{\mathbf{N}''})+k-1$$

A jobb oldalon (2.8) és (2.9) miatt $s(\mathscr{L}_{\mathbf{N}})+k-1$ is írható, azaz

$$\mathbf{N} \text{ sorainak száma } \geqq s(\mathscr{L}_N)+k,$$

amint azt bizonyítani kívántuk.

Térjünk vissza a tétel, pontosabban a $\geqq$ egyenlőtlenség bizonyítására. Legyen tehát valamely $\mathbf{M}$-re $\mathscr{L}_{\mathbf{M}}=\mathscr{L}_1\times\mathscr{L}_2$ és az $\mathscr{L}_1$ és $\mathscr{L}_2$ alaphalmazának megfelelő részmátrixok $\mathbf{M}_1$ és $\mathbf{M}_2$. Az első lemma szerint $\mathscr{L}_{\mathbf{M}_1}=\mathscr{L}_1$, de ennél valamivel több is igaz. Ha $a\notin\mathscr{L}_1(A)$ $(A\subset X_1)$, akkor $a\notin\mathscr{L}_1(A)\cup X_2=(\mathscr{L}_1\times\mathscr{L}_2)$ $(A\cup X_2)$, azaz $\mathbf{M}$-nek van két olyan sora, melyek $A\cup X_2$-ben megegyeznek, $a$-ban nem. Azaz, $\mathbf{M}_1$-ben két olyan sor található, 1. melyekhez tartozó $\mathbf{M}_2$-beli rész megegyezik, 2. $A$-ban megegyeznek, 3. $a$-ban nem egyeznek meg. Vagyis $\mathbf{M}_1$ sorait ($\mathbf{M}_2$ különböző sorainak száma) számú olyan osztályba soroltuk, melyek kielégítik a második lemma feltételeit. Így (2.7) alapján

$$\mathbf{M}_1 \text{ sorainak száma } \geqq s(\mathscr{L}_1)+(\mathbf{M}_2 \text{ különböző sorainak száma})-1.$$

Az első lemma alapján $\mathbf{M}_2$ különböző sorainak száma $\geqq s(\mathscr{L}_2)$. Innen a kívánt egyenlőtlenség következik.

## 3. Kulcsrendszerek direkt szorzata

Ha $\mathscr{K}_1$, illetve $\mathscr{K}_2$ a minimális kulcsok rendszere a diszjunkt $X_1$, illetve $X_2$ halmazokon, akkor $\mathscr{K}_1\times\mathscr{K}_2$ álljon $X_1\cup X_2$ azon halmazaiból, amelyek $A\neq B$, $A\in\mathscr{K}_1$, $B\in\mathscr{K}_2$ alakban állnak elő. Könnyen látható, hogy $\mathscr{K}_1\times\mathscr{K}_2$ egy *Sperner-rendszer,* azaz van olyan lezárás, amelyben ezek a minimális kulcsok.

3.1. TÉTEL.

$$s(\mathcal{K}_1\times\mathcal{K}_2) \leqq s(\mathcal{K}_1)+s(\mathcal{K}_2)-1. \tag{3.1}$$

*Bizonyítás.* Az egyenlőtlenség a 2.1. tétel bizonyításának első részével azonos módon bizonyítható. Az ott leírt konstrukció itt is megfelel. A tapasztalat azt mutatja, hogy (3.1)-ben csaknem mindig egyenlőtlenség áll, a 2.1. tétel esetéhez hasonóan. Azonban a következő ellenpélda mutatja, hogy ez nem mindig igaz.

Legyen

$$X_1 = \{1, 2, 3, 4, 5\},\quad X_2 = \{6, 7, 8, 9, 10\},$$

$$\mathcal{K}_1 = \{\{1, 2\}, \{3\ 4\}, \{1, 5\}, \{2, 5\}, \{3, 5\}, \{4, 5\}\}$$

és $\mathcal{K}_2$ ezzel izomorf $X_2$-n. A maximális nem kulcsok rendszere $X_1$-en $\mathcal{K}^{-1} = \{\{5\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$.

Lássuk be először, hogy $s(\mathcal{K}_1)\geqq 5$. Tegyük fel indirekte, hogy van egy 4 soros $\mathbf{M}_1$ mátrix, amelynek kulcsrendszere $\mathcal{K}_1$. Fel fogjuk használni a következő egyszer lemmát:

3.1. LEMMA. Ha $\mathbf{M}$ kulcsrendszere $\mathcal{K}$, akkor $\mathcal{K}^{-1}$ minden $A$ eleméhez létezik egy sorpár $\mathbf{M}$-ben, melyek pontosan $A$-ban egyeznek, és $\mathcal{K}^{-1}$ különböző elemeihez különböző sorpárok tartoznak.

*Megjegyzés:* A lemma állítása implicite megjelenik már [1]-ben is.

*Bizonyítás.* Mivel minden $A$ nem-kulcshoz tartozik egy sorpár, melyek egyenlők $A$-ban, de nem teljesen egyenlők, így $A\in\mathcal{K}^{-1}$ esetén is van ilyen. Ha ez a két sor több helyen is egyenlő lenne nem csak $A$-ban, akkor $A$-nak volna egy olyan valódi bővítése, ami $\mathcal{K}^{-1}$-ben volna, ellentmondásban $A\in \mathcal{K}^{-1}$-gyel.

Ebből máris következik, hogy $s(\mathcal{K}_1)\geqq 4$, hiszen $\binom{3}{2}=3<5=|\mathcal{K}_1^{-1}|$. De ha $\mathbf{M}$ sorainak száma 4, akkor is egy kivételével minden sorpár egyezési halmaza valamelyik $\mathcal{K}_1^{-1}$-beli elem. Még egy egyszerű lemmát használunk:

3.2. LEMMA. Egy mátrix 3 sorának 3 egyezési halmazára $A_1, A_2, A_3$-ra fennáll

$$(A_i\cap A_j)-A_k = \emptyset, \tag{3.2}$$

ahol $i, j, k$ az $1, 2, 3$ egy permutációja.

*Bizonyítás.* Triviális.

Az $\mathbf{M}$ 4 sorából képezhető 6 párból 5-höz tartozik $\mathcal{K}_1^{-1}$-beli halmaz. (Legyen a kivétel a $(3, 4)$ sorpár.) Ez az 5 pár 2 darab sorhármast határoz meg. Ezekre teljesülni kell (3.2)-nek. Mivel $\mathcal{K}_1^{-1}$-ben nincs 3 halmaz, amelynek a metszete nem üres, így a 3 halmaznak páronként is diszjunktnak kell lennie. 3 halmaz (szimmetriától eltekintve) csak egyféleképpen választható így ki: $\{1, 3\}$, $\{2, 4\}$, $\{5\}$. De ekkor a másik hármas egyértelmű: $\{1, 4\}, \{2, 3\}, \{5\}$, ugyanis a két hármasnak nem lehet egynél több közös halmaza. Szimmetriaokok miatt feltehető tehát, hogy az 1. és 2. sor megegyezési halmaza $\{5\}$, az 1. és 3.-é $\{1, 3\}$, a 2. és 3.-é $\{2, 4\}$, továbbá az 1. és 4.-é $\{2, 3\}$. De ekkor a 3. és 4. sor megegyezik $\{1, 2\}$-ben, de $\{3, 4\}$-ben nem, azaz $\mathcal{K}_1^{-1}$-nek kellene lennie egy $\{1, 2\}$-t tartalmazó elemének. Ez az ellentmondás bizonyítja, hogy $s(\mathcal{K}_1)\geqq 5$.

Ha (3.1)-ben egyenlőség állna, $s(\mathcal{K}_1 \times \mathcal{K}_1) = s(\mathcal{K}_1^2) = 9$ lenne. Egy konstrukcióval igazolni fogjuk, hogy $s(\mathcal{K}_1^2) \leqq 8$.

Használjuk a következő jelöléseket:

$$\alpha = 0,\, 0,\, 0,\, 0,\, 0$$

$$\beta = 1,\, 1,\, 1,\, 1,\, 0$$

$$\gamma = 0,\, 1,\, 0,\, 1,\, 1$$

$$\delta = 1,\, 0,\, 0,\, 1,\, 1$$

A mátrix, amelyik $s(\mathcal{K}_1^2) \leqq 8$-at bizonyítja, a következő:

$$\mathbf{M} = \begin{matrix} \alpha & \alpha \\ \alpha & \beta \\ \alpha & \gamma \\ \beta & \alpha \\ \beta & \beta \\ \beta & \delta \\ \gamma & \alpha \\ \delta & \beta \end{matrix}$$

Lássuk be, hogy $\mathbf{M}$ minimális kulcsainak halmaza éppen $\mathcal{K}_1^2$. Ehhez először azt kell igazolnunk, hogy a $\mathcal{K}_1^2$-beli halmazok $\mathbf{M}$-ben kulcsok, azaz nincs két olyan sora $\mathbf{M}$-nek, melyek különbözők, de egy $\mathcal{K}_1^2$-beliben vagy egy ilyennek bővítésében egyeznek meg. Vagyis két sor vagy bal vagy jobb oldalán nem $\mathcal{K}_1$-beli halmazban vagy ilyennek bővítésében egyezik. Könnyen ellenőrizhető, hogy $\alpha, \beta, \gamma, \delta$ közül csak egy pár, $\gamma$ és $\delta$ egyezik $\mathcal{K}_1$-beliben vagy bővítésében. Természetesen az egyformák is ilyenek. A két oldal összehasonlítása mutatja, hogy valóban nincs két olyan teljes sor, mely mindkét oldalon $\mathcal{K}_1$-beliben vagy bővítésében egyezne.

Másrészt azt is igazolni kell, hogy $(\mathcal{K}_1^2)^{-1}$ elemei nem kulcsok. Könnyen látható, hogy $(\mathcal{K}_1^2)^{-1}$ az olyan halmazokból áll, melyek egyik oldalon egy $\mathcal{K}_1^1$-beliből, a másik oldalon a teljes alaphalmazból állnak. Az ilyen halmazokra kell találni sorpárt $\mathbf{M}$-ben, melyek itt megegyeznek, de nem azonosak. A két rész szimmetriája miatt feltehető, hogy a keresendő két sor az első félben egyezik. Könnyen áttekinthető, hogy ekkor a második félben minden $\mathcal{K}_1^{-1}$-belihez található két olyan sor, ami ebben egyezik, hozzá tartozó első fele pedig egyforma. Ezzel a konstrukció összes szükséges tulajdonságát igazoltuk.

## 4. Telített Sperner-rendszerek és lezárások

Egy *Sperner-rendszert* akkor nevezünk *telített*nek, ha a rendszer nem bővíthető további halmazzal úgy, hogy a *Sperner-tulajdonsága* megmaradjon. Ha $\mathcal{L}$ egy lezárás az $X$ alaphalmazon, a

$$\{K\colon \mathcal{L}(K) = X,\ \ \nexists K' \subset K,\ \ \mathcal{L}(K') = X\}$$

rendszert, $\mathcal{L}$ minimális kulcsainak rendszerét $\mathcal{K}(\mathcal{L})$-lel jelöljük.

4.1. TÉTEL. Ha $\mathscr{K}$ egy telített *Sperner-rendszer*, akkor a $\mathscr{K}(\mathscr{L})=\mathscr{K}$ egyen-lőség egyértelműen meghatározza az $\mathscr{K}$ lezárást.

*Bizonyítás*. Igazoljuk, hogy ha $\mathscr{K}$ telített és $\mathscr{K}(\mathscr{L})=\mathscr{K}$, akkor minden $A$ esetén meg van határozva. $\mathscr{K}$ telítettsége miatt minden $A$ halmazhoz van egy $K\in\mathscr{K}$, amelyre 1. $A\supseteqq K$ vagy 2. $A\subset K$, $A\neq K$ teljesül. Az első esetben $\mathscr{L}(A)=X$, tehát a lezárás valóban egyértelmű. A második esetben azt fogjuk belátni, hogy $\mathscr{L}(A)=A$. Tegyük fel indirekte, hogy $a\in\mathscr{L}(A)-A$. A telítettség miatt ismét van egy $K_1\in\mathscr{K}$ halmaz, amire vagy 2/a) $A\cup\{a\}\subseteqq K_1$ vagy 2/b) $A\cup\{a\}\supset K_1$, $A\cup\{a\}\neq$ $\neq K_1$ teljesül. A 2/a) esetben (2.2) alapján $\mathscr{L}(K_1-\{a\})\supseteqq K_1-\{a\}$, (2.3) alapján (mivel $K_1-\{a\}\supseteqq A)/\mathscr{L}(K_1-\{a\})\supseteqq\mathscr{L}(A)\supseteqq A\cup\{a\}$ állítható. Összegezve $\mathscr{L}(K_1-$ $-\{a\})\supseteqq K_1$. Amiből (2.3a)-val $\mathscr{L}(K_1-\{a\})=\mathscr{L}((K_1-\{a\}))\supseteqq\mathscr{L}(K_1)=X$ követ-kezne, ellentmondásban $K_1$ minimalitásával. A 2/b) esetben $\mathscr{L}(A)\supseteqq A\cup\{a\}\supseteqq K_1$, vagyis $\mathscr{L}(A)=\mathscr{L}(\mathscr{L}(A))\supseteqq\mathscr{L}(K_1)=X$, ismét ellentmondással. Tehát egy $\mathscr{K}$ *telített Sperner-rendszer* egyértelműen határozza meg $\mathscr{L}$-et.

A tétel megfordítása azonban nem igaz, mint azt a következő példa mutatja. Legyen az alaphalmaz, $X$ 5-elemű és $\mathscr{K}=\{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,5\}\}$. Iga-zoljuk, hogy — bár $\mathscr{K}$ nem telített — egyértelműen meghatározza minden halmaz lezárását. Minden 2-nél több elemű halmaz tartalmaz egy $\mathscr{K}$-beli halmazt, ezek lezá-rása tehát $X$. A $\mathscr{K}$-beli 2-eleműek lezárása $X$. A $\mathscr{K}$-ban nem szereplő 2-eleműek lezárása önmaga. Ugyanis, ha pl. $\mathscr{L}(\{1,3\})$ legalább 3-elemű volna, akkor tartal-mazna egy $\mathscr{K}$-belit, így $\mathscr{L}(\mathscr{L}(\{1,3\}))=\mathscr{L}(\{1,3\})$ tartalmazná $X$-et, azaz $\mathscr{L}(\{1,3\})=$ $=X$ lenne. Mivel sem $\{1,3\}$, sem valamilyen része nincs $\mathscr{K}$-ban, ellentmondásra jutottunk. Ha most $\mathscr{L}(\{1\})$-nek volna 1-től különböző eleme, akkor azt $\mathscr{L}(\{1,3\})$ és $\mathscr{L}(\{1,4\})$ is tartalmazná. Egyikük ellentmondást okoz. Hasonlóan látható, hogy minden egyelemű halmaz lezárása önmaga, és $\mathscr{L}(\{\emptyset\})=\emptyset$.

Végtelen sok ilyen példát találtunk, de nem sikerült pontosan jellemeznünk azon $\mathscr{K}$-kat, amelyek egyértelműen meghatároznak egy lezárást[1].

## IRODALOM


[1] DEMETROVICS, J. and KATONA, G. O. H., "Extremal combinatorical problems in relational date base", *Fundamentals of Computation Th., Lecture Notes in Computer Science* **117** (Springer Verlag, Berlin, Heidelberg, N. Y., 1981) 110—119.

[2] DEMETROVICS, J. és KATONA, GY., „A relációs adatbázis extremális problémái", *Alkalmazott Matematikai Lapok* **8** (1982) 183—193.

[3] CODD, E. F., "A relational model of data for large shared data banks", *Comm. ACM* **13** (1970) 377—387.

[4] DEMETROVICS, J., "On the equivalence of candidate keys with Sperner systems", *Acta Cybernetica* **4** (1979) 247—252.

[5] DEMETROVICS, J. and GYEPESI, GY., "On the functional dependency and some generalization of it", *Acta Cybernetica.*


(Beérkezett: 1982. december 15.)


DEMETROVICS JÁNOS
MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1132 BUDAPEST, VICTOR HUGO U. 18.

FÜREDI ZOLTÁN ÉS KATONA GYULA
MTA MATEMATIKAI KUTATÓ INTÉZET
1053 BUDAPEST, REÁLTANODA U. 13—15.


[1] *Megjegyzés a korrektúránál.* Azóta BUROSCH G.-vel közsen sikerült jellemeznünk azon $\mathscr{K}$-kat, melyek egyértelműen meghatároznak egy lezárást.

# THE RELATION BETWEEN THE NUMBER OF INDIVIDUALS AND DEPENDENCIES IN THE COMPOSED DATA BASES

J. DEMETROVICS, Z. FÜREDI and G. O. H. KATONA

In an earlier paper we have investigated the problem that how many is the minimal number of individuals in a data base with a given key system. In the present paper we investigate the same question and a similar one. Namely we shall determine the minimal number of individuals in that case when in the data base a system of dependencies is also given.

In many cases it is very useful to determine some parameters of a composed data base from the same parameters of the sub data bases. We try to do this for the minimal number of the individuals. We can do this only if the data base is composed from independent parts.