Alkalmazott Matematikai Lapok **24** (2007), 223-264.

RELÁCIÓS ADATBÁZISOK FUNKCIONÁLIS FÜGGŐSÉGI RENDSZEREINEK GRAFIKUS AXIOMATIZÁCIÓJA[1]

DEMETROVICS JÁNOS, MOLNÁR ANDRÁS ÉS BERNHARD THALHEIM

Relációs adatbázisok tervezésekor az adattáblák sémái mellett fontos azt is megadni, hogy az adatbázisnak milyen invariáns tulajdonságokat kell teljesítenie. E feltételeket integritási megszorítások formájában adhatjuk meg, melyek közül a funkcionális függőségek a legalapvetőbbek. A tervezési gyakorlatban azonban a formalizmus használata nehézkes, így gyakran csak közvetve alkalmazzák. Egyes esetekben viszont az egyértelmű és teljes specifikáció csak függőségi rendszer megadásával lehetséges. Ehhez egy olyan diagramos reprezentációt készítettünk, amely lehetővé teszi a függőségi rendszerek hagyományos jelölésnél egyszerűbb, áttekinthetőbb kezelését. A grafikus reprezentáció a hozzá kifejlesztett axiomatizációval az implikált függőségek levezetését természetes módon támogatja, kisebb sémákra pedig kézzel is könnyen elvégezhetővé teszi. Ezen kívül kis sémákra (legfeljebb 5 attribútum esetén) a lehetséges függőségi rendszerek, ill. típusok számát is meghatároztuk.

1. Bevezetés

Egy adatbázis struktúrájának tervezésekor az alábbi három fő szempont játszik szerepet [27]. Ezek természetesen nem függetlenek egymástól.

Szintaktika: Az adatbázis logikai szerkezetének megadása induktív módon, alaptípusokból kiindulva, konstrukciós műveletek segítségével. A műveletek alkalmazhatósági feltételei logikai szabályokkal írhatók le.

Szemantika: Az adatbázis elfogadható állapotainak leírása. A megadott logikai feltételeket az adatbázis invariánsan kell tartsa, ellenkező esetben az adatok nem lennének értelmezhetők.

Pragmatika: Az adatbázis környezetének és a környezetben betöltött szerepének megadása. Alapja a tárolt adatok és az elvégezhető műveletek kapcsolata az adott intézménnyel vagy céggel, az intézményi, ill. üzleti munkafolyamatokkal, a felhasználókkal.

A szintaktikai rész egyszerű esetben megadható közvetlenül a használt adatmodellben (pl. relációs modell [16], ld. később e fejezetben), de általában valamely

[1]Készült az OTKA támogatásával (T042706).

modellező nyelven tervezik meg az adatbázis szerkezetét (a nyelv többnyire vizuális, pl. egyed-kapcsolat modell [14]), s az elkészült tervet alakítják át az adatmodell sémájára. A pragmatikai rész megadása gyakran nem teljesen explicit. A szemantikai részben az adatbázis-invariánsok itegritási megszorítások (függőségek) formájában adhatók meg (pl. egy azonosító szám valóban egyetlen egyedhez tartozzon, ne fordulhasson elő többször). Ezek sokfélék lehetnek; az egyes típusokhoz formális logikai nyelvek állnak rendelkezésre; többségük axiomatizálható [1]. A szemantikai rész megadása sokszor elbonyolódik, mert nehéz áttekinteni a függőségek közti kapcsolatokat és végiggondolni a lehetséges függőségek érvényességét. Ehhez szükség van módszerekre, melyek segítik az adatbázis tervezőjét meghatározni az eleve érvényes és érvénytelen függőségeket és azok logikai következményeit. Bár a modellező nyelvek többsége ad eszközöket ilyen megszorítások megadására, ez egyes esetekben a teljesen pontos specifikációra nem elegendő; logikai következtetést például általában nem tesznek lehetővé.

1.1. A relációs adatmodell

A relációs modell [16] máig is a legszélesebb körben alkalmazott adatmodell, elsősorban egyszerűsége, kifejezőereje és hatékony implementálhatósága révén [1, 32]. Az adatbázis előre meghatározott típusú adattáblákból áll; egy tábla sorok halmaza; egy sor pedig a tábla típusában (sémájában) meghatározott attribútumokhoz rendel értékeket. Ez pontosabban az alábbiakat jelenti.[2]

Legyen $Attr$ és Dom szimbólumok egy-egy megszámlálható halmaza, melyekre értelmezett a $DOM : Attr \to 2^{Dom}$ függvény. $Attr$ az attribútumszimbólumok, Dom pedig az adatbázisban előforduló szimbólumok (elemi adatok) univerzuma. DOM adja meg az attribútumok típusát (értéktartományát).

Legyen $\mathcal{R} \subseteq Attr$ attribútumszimbólumok egy véges, nem üres halmaza (ún. *relációséma*), R pedig $r : \mathcal{R} \to Dom$ alakú függvények véges halmaza. R-et $\mathcal{R}$ sémájú relációnak (*reláció-előfordulásnak*) nevezzük, ha

$$\forall r \in R : \forall A \in \mathcal{R} : r(A) \in DOM(A).$$

R attribútumai (vagy *oszlopai*) alatt $\mathcal{R}$ elemeit értjük, *R sorai* alatt pedig magának az R relációnak az elemeit.

Legyen Rel egy megszámlálható szimbólumhalmaz, a *relációséma-szimbólumok* halmaza és T egy véges részhalmaza. Egy $\mathcal{DB} : T \to 2^{Attr}$ függvényt *relációs adatbázissémának* nevezünk; $\mathcal{DB}$ sémájú *relációs adatbázisnak* pedig egy olyan DB függvényt, amely T minden t elemét egy $\mathcal{DB}(t)$ sémájú relációra képezi le. Ezek a relációk az adatbázis táblái, T elemei a táblák nevei.

Például egy napi repülőgép-járatokat nyilvántartó J adattábla relációsémája lehet

$$\mathcal{J} = \{Járatszám, Honnan, Hová, Mikor\},$$

[2] Többféle, lényegében ekvivalens definíció is adható [1]. Mi az attribútumokat típusosnak tekintjük, relációsémabeli sorrendjüket tetszőlegesnek vesszük, az attribútumokat névvel azonosítjuk. Így a definíciót funkcionális alapon adjuk meg.

melynek az egyik eleme (sora) lehet

$$r = \{Járatszám \mapsto 6209,\ Honnan \mapsto Budapest,\\ Hová \mapsto London,\ Mikor \mapsto 18:05\}.$$

Ha rögzítjük az attribútumok sorrendjét, a jelölés egyszerűsödik:

$$r = (6209,\ Budapest,\ London,\ 18:05).$$

Ha nyilvántartjuk a pilótákat és azt is, hogy melyik járatot ki vezeti közülük, akkor az előbbiek bővítéseként egy három táblából álló adatbázissémánk lehet, $T = \{\mathcal{P}, \mathcal{J}, \mathcal{V}\}$ táblanevekkel. Itt $\mathcal{P}$ a pilóták adattáblájának sémája, $\mathcal{V}$ („vezeti") pedig egy olyan séma, amely a pilóták és a járatok közti hozzárendelésnek felel meg, előfordulásának sorai egy-egy pilóta azonosítóját és járatszámot tartalmaznak.

1.1.1. Egyed-kapcsolat diagram

A relációs adatbázissémák tervezésének legelterjedtebb modellező nyelve az *egyed-kapcsolat diagram* (pl. [14, 15, 13, 30, 32]). A nyelv nem egységes, többféle változata létezik.

A tervezés során tekintjük a valós világnak azon részét, amelyről adatokat kívánunk tárolni és *egyedosztályokat* különítünk el, *attribútumokkal* (a fenti példában egyedosztályként modellezhetjük a járatokat és a pilótákat). Az egyedosztályok közötti viszonyokat pedig *kapcsolatok* formájában modellezzük (pl. a pilóták és a járatok között kapcsolatként jelenik meg, hogy ki melyik járatot vezeti). A kapcsolatok sokfélék lehetnek, mind aritásukat (pl. bináris – két egyedosztály közötti; ternáris; kvaternáris), mind pedig a részt vevő egyedek multiplicitását tekintve (pl. egy járatot hány pilótának kell vezetnie, egy pilóta hány járatot vezethet). A bináris kapcsolatoknak három alaptípusa van: egy-egy (pl. a pilóták és a járatok között egy-egyértelmű a megfeleltetés[3]), egy-sok (egy járathoz egy pilóta tartozik, de egy pilóta többet is vezethet, azaz fordított irányban nincs függvénykapcsolat), sok-sok (egyik irányban sem adható függvénykapcsolat). Mi az olyan típusú kapcsolatokkal foglalkozunk, amelyek ilyen jellegű viszonyokkal megfogalmazhatók.

Az adatmodell diagram formájában készül el, mely átalakítható relációs adatbázissémává: az egyedosztályokból relációsémák lesznek, általában a kapcsolatokból is („kapcsolótábla", pl. a fenti $\mathcal{V}$), de a séma gyakorta egyszerűsíthető.

1.2. Funkcionális függőségek és negáltjaik

1.2.1. Funkcionális függőségek

Az integritási megszorítások legalapvetőbb típusa a *funkcionális függőség* [2], melynek elmélete jól ismert és kidolgozott, számos fontos fogalom épül rá (kulcsok,

[3]Tágabb értelemben a kapcsolatban való részvétel lehet opcionális, azaz a pilótáknak és a járatoknak csak egy-egy részhalmazán értelmezett a bijekció.

dekompozíciók, normálformák; pl. [6, 7, 5, 8, 31, 32]). Egy funkcionális függőség azt jelenti, hogy a relációs adattáblában egyes attribútumok értékei másokkal függvénykapcsolatban kell legyenek, azaz egyértelműen meghatározzák azokat. Pontosabban:

Legyen $\mathcal{R}$ egy relációséma, X és Y pedig attribútumainak egy-egy részhalmaza. Ekkor $X \to Y$ egy *funkcionális függőség*, melyet egy $\mathcal{R}$ sémájú R reláció akkor és csak akkor *teljesít* (jel.: $R \vDash X \to Y$), ha

$$\forall r, s \in R : (\forall A \in X : r(A) = s(A)) \Rightarrow (\forall B \in Y : r(B) = s(B))$$

(egyszerűbben jelölve $r(X) = s(X) \Rightarrow r(Y) = s(Y)$), azaz ha két sor X-en megegyezik, akkor Y-on is. A fogalom értelemszerűen általánosítható függőségek halmazaira is.

Jelöléseinkben két attribútumhalmaz unióját $X \cup Y$ helyett XY-nal jelöljük a szokásos módon; s ahol nem okoz zavart, ott az egyetlen attribútumból álló halmazt ($\{A\}$) maga az attribútum (A) jelöli.

Kanonikus (vagy szingleton) függőségnek nevezzük azt, amelynek a jobb oldalán egyetlen attribútum áll; *triviálisnak* pedig azokat tekintjük, melyeknél a jobb oldal része a bal oldalnak. *Redundánsak* azok a függőségek (ezeket lehetne akár *pszeudotriviálisnak* is nevezni), melyeknek van bal és jobb oldalukon közös attribútumuk (vagy a jobb oldaluk üres). Ezek érvényessége ekvivalens azokéval, melyeket a jobb oldalból a bal oldalon is szereplő rész kivonásával kapunk, ezért a redundáns függőségek kezelését (köztük a triviálisokét) igyekszünk majd elkerülni.

A funkcionális függőségek körében értelmezhető a *logikai következmény* fogalma. Ha ζ egy függőség és $\mathcal{F}$ függőségek egy (véges) halmaza, akkor ζ pontosan akkor logikai következménye $\mathcal{F}$-nek (jel.: $\mathcal{F} \vDash \zeta$), ha minden R relációra

$$R \vDash \mathcal{F} \Rightarrow R \vDash \zeta.$$

Ha $\mathcal{F}$ funkcionális függőségek egy halmaza, akkor $\mathcal{F}$ *lezártja* alatt azt az $\mathcal{F}^+$-szal jelölt halmazt értjük, mely $\mathcal{F}$ logikai következményeit tartalmazza:

$$\mathcal{F}^+ = \{\delta = X \to Y \mid \mathcal{F} \vDash \delta\}.$$

$\mathcal{F}$ *zárt*, ha $\mathcal{F} = \mathcal{F}^+$.

1.2.2. Negált funkcionális függőségek

Egy adatbázis tervezésekor a valós világot modellezzük, s egyes jelenségek funkcionális függőségek formájában kerülnek be a modellbe (erre még később visszatérünk). A tervezés közben a függőségi rendszert finomítjuk, azaz egyre több függőséget veszünk hozzá a halmazhoz. Így minden lépésben nyitott kérdés, hogy azok a lehetséges függőségek, melyek még nem szerepelnek a halmazban, vajon igazak-e vagy sem. Egy adott lehetséges függőség vizsgálatakor kétféle módon derülhet ki, hogy be kell vennünk a modellbe: a vizsgált függőség logikai következménye az eddigieknek, ezért igaz; vagy a vizsgált függőséget a valós világgal összevetve azt találjuk, hogy szükségképpen igaznak kell deklarálnunk a modellben.

A vizsgálat viszont azt is eredményezheti, hogy a függőség nem szükségképpen igaz, azaz a majdani adatbázisban meg kell engednünk, hogy a függőségre ellenpélda legyen: két olyan sor, amely a bal oldalnak megfelelő attribútumokon megegyezik, de a jobb oldalnak megfelelőeken különbözik. Ennek figyelembe vételéhez a funkcionális függőségek mellett ún. *negált függőségeket* is kezelünk, melyek $X \nrightarrow Y$ alakúak. Egy ilyen nem integritási megszorításként értelmezendő, hiszen nem írja elő, hogy az adatbázisnak csak olyan állapotai lehetnek, melyben találunk ellenpéldát az $X \to Y$ függőségre. Ehelyett a tervezési fázisban van szerepe, és azt fejezi ki, hogy nem írhatjuk elő az $X \to Y$ függőséget, azaz a tervezés végén előállítandó funkcionális függőségi rendszernek csak olyan halmaz fogadható el, melyben az adott függőség nem szerepel. A fent bevezetett *kanonikus (szingleton)* és *redundáns* fogalmakat értelemszerűen általánosíthatjuk a negált függőségekre.

A funkcionális függőségeket ezzel összefüggésben gyakran *pozitív* függőségeknek nevezzük, míg a negált függőségeket *negatív*aknak. Egy rögzített megszámlálható attribútumszimbólum-készlet felett kifejezhető összes funkcionális függőség univerzumát jelöljük $\mathbb{D}$-vel, a negált függőségek univerzumát pedig $\mathbb{E}$-vel. Egy adott ζ pozitív (vagy negatív) függőség *negáltja* alatt a neki megfelelő negatív (ill. pozitív) függőséget értjük (azaz amelynek jobb és baloldala megegyezik ζ-val, de ellenkező értelmű), és $\neg\zeta$-val jelöljük. Ezt a jelölést értelemszerűen általánosíthatjuk függőségek halmazaira. Egy $\mathcal{F} \subseteq \mathbb{D} \cup \mathbb{E}$ függőségi halmaz egy adott attribútumhalmazra nézve *teljes*, ha nem tartalmaz egyszerre egy függőséget és annak negáltját, továbbá a $\mathcal{F} \cap \mathbb{D} \cup \neg(\mathcal{F} \cap \mathbb{E})$ halmaz éppen az adott attribútumhalmaz felett felírható összes lehetséges funkcionális függőséget tartalmazza (azaz minden lehetséges függőségre a halmaz megadja, hogy a modellben igaz-e vagy sem).

Egy $\mathcal{F} \subset \mathbb{D} \cup \mathbb{E}$ függőségi halmaz *lehetséges kiterjesztése* minden olyan $\mathcal{G}^+ \subset \mathbb{D}$ zárt funkcionális függőségi halmaz, melyre

$$\mathcal{F} \cap \mathbb{D} \subseteq \mathcal{G}^+ \quad \text{és} \quad \neg(\mathcal{F} \cap \mathbb{E}) \cap \mathcal{G}^+ = \emptyset$$

(azaz tartalmazza $\mathcal{F}$ pozitív függőségeit, de nem tartalmazza $\mathcal{F}$ negatív függőségeinek negáltjait). A definícióban szereplő $\mathcal{G}^+$ halmazok valójában azok a lehetséges függőségi rendszerek, melyek $\mathcal{F}$ finomításával (további függőségek hozzávételével, teljessé tétellel) megkaphatók, azaz a tervezés adott fázisában lehetséges kimeneteleknek tekinthetők.

Pozitív és negatív függőségek együttes kezelésekor a logikai következmény fogalmát másképpen kell értenünk. Tegyük fel, hogy $\mathcal{F} \subset \mathbb{D} \cup \mathbb{E}$ egy (véges) függőségi halmaz, és $\zeta \in \mathbb{D} \cup \mathbb{E}$ egy függőség. Ekkor ζ pontosan akkor *logikai következménye* $\mathcal{F}$-nek, ha minden lehetséges $\mathcal{G}^+ \subset \mathbb{D}$ kiterjesztésére teljesül, hogy $\zeta \in \mathcal{G}^+$, amennyiben ζ pozitív; illetve $\neg\zeta \notin \mathcal{G}^+$, amennyiben ζ negatív. Könnyű látni, hogy ha ζ pozitív, és $\mathcal{F} \subset \mathbb{D}$, akkor ez éppen ugyanazt jelenti, mint a korábban definiált következményfogalom. A *lezárás* fogalma újraértelmezhető ezzel a logikai következményfogalommal.

Egy $\mathcal{F}$ függőségi halmaz *ellentmondásos*, ha nem létezik lehetséges $\mathcal{G}^+$ kiterjesztése. Ekkor $\mathcal{F}$-nek logikai következménye bármilyen függőség, így egyidejűleg valamely $X \not\twoheadrightarrow Y$ és $X \to Y$ függőség is. Ennek vizsgálata, felismerése fontos a tervezés során, hiszen ilyenkor valami hibát vétettünk a modellezésben és finomítással nem tudunk eljutni érvényes, teljes függőségi rendszerhez.

1.2.3. Afunkcionális függőségek

Az imént bevezetett negált függőségfogalomnál létezik egy erősebb szemantikájú is, mely azt fejezi ki, hogy az adatbázisban mindig lennie kell olyan soroknak, melyek egy adott funkcionális függőséget megsértenek. Az ilyet *afunkcionális* függőségnek nevezzük [9].

Az afunkcionális függőségek számunkra kevésbé érdekesek, de az alábbi összefüggés rámutat arra, hogy formálisan azonosan kezelhetők a mi negált függőségeinkkel, hiszen a logikai következményfogalom a két esetben ekvivalens:

1.1. Állítás. *Legyen $\mathcal{F} \subset \mathbb{D} \cup \mathbb{E}$ egy véges halmaz és $\zeta \in \mathbb{E}$ egy negált függőség. Ekkor $\mathcal{F} \models \zeta$ ekvivalens a következővel: minden olyan R relációban, amelyre $R \models \mathcal{F} \cap \mathbb{D}$, de $\forall \delta \in \mathcal{F} \cap \mathbb{E} : R \not\models \neg\delta$, igaz az is, hogy $R \not\models \neg\zeta$ (azaz ha R kielégíti $\mathcal{F}$ pozitív függőségeit, és van benne ellenpélda $\mathcal{F}$ negatív függőségeinek negáltjaira – így a nekik megfelelő afunkcionális függőségeket R kielégíti –, akkor kell lennie ellenpéldának $\neg\zeta$-ra is – s így a ζ-nak megfelelő afunkcionális függőség teljesül).*

Bizonyítás. Az ekvivalencia egyik irányának igazolásához tegyük fel, hogy $\mathcal{F} \models \zeta$, és R egy olyan reláció, amely $\mathcal{F}$ minden pozitív függőségét kielégíti, de minden negatív függőségének negáltjára található R-ben ellenpélda (cáfoló két sor). Indirekt módon tegyük fel, hogy $R \models \neg\zeta$, azaz nincs ellenpélda $\neg\zeta$-ra. Ekkor tekintsük azt a $\mathcal{G}^+ \subset \mathbb{D}$ függőségi rendszert, mely pontosan azokat a funkcionális függőségeket tartalmazza, melyeket R kielégít. $R \models \neg\zeta$, ezért $\neg\zeta \in \mathcal{G}^+$. Ugyanakkor ez a $\mathcal{G}^+$ zárt, és R választása miatt $\mathcal{F}$ egy lehetséges kiterjesztése, így ellentmondásra jutunk azzal, hogy $\mathcal{F} \models \zeta$.

A másik irányhoz vegyünk egy olyan $\mathcal{G}^+ \subset \mathbb{D}$ zárt funkcionális függőségi rendszert, mely $\mathcal{F}$ egy lehetséges kiterjesztése. Ahhoz, hogy megmutassuk $\mathcal{F} \models \zeta$-t, meg kell mutatnunk, hogy $\neg\zeta \notin \mathcal{G}^+$. Ehhez vegyük a $\mathcal{G}^+$ egy R Armstrong-relációját [26, 4, **?**], azaz egy olyan relációt, amely éppen $\mathcal{G}^+$ függőségeit elégíti ki. Mivel R attribútumain felírható minden más funkcionális függőségre lesz ellenpélda R-ben, így a $\mathcal{F}$-ben szereplő negatív függőségek negáltjaira is, a feltétel (a bizonyítandó ekvivalencia másik oldala) szerint lesz ellenpélda $\neg\zeta$-ra is. Emiatt $R \not\models \neg\zeta$, amely éppen azt jelenti, hogy $\neg\zeta \notin \mathcal{G}^+$, mivel R Armstrong-reláció. □

1.2.4. A függőségek hagyományos axiomatizációja

A funkcionális függőségek ismert helyes és teljes következtetési szabályrendszere az alábbi, ún. *Armstrong-axiómarendszer* [2]:

$$(0) \quad XY \to Y$$

$$(1) \quad \frac{X \longrightarrow Y}{XVW \longrightarrow YV} \qquad (2) \quad \frac{X \longrightarrow Y,\ Y \longrightarrow Z}{X \longrightarrow Z}$$

Negált függőségek esetén az axiomatizációt az alábbi levezetési szabályokkal bővítve kapjuk a *kiterjesztett Armstrong-axiómarendszert*[4], mely úgyszintén helyes és teljes [30]:

$$(3) \quad \frac{X \longrightarrow Y,\ X \not\longrightarrow Z}{Y \not\longrightarrow Z} \qquad (4) \quad \frac{X \not\longrightarrow Y}{X \not\longrightarrow YZ} \qquad (5) \quad \frac{XZ \not\longrightarrow YZ}{XZ \not\longrightarrow Y}$$

$$(6) \quad \frac{X \longrightarrow Z,\ X \not\longrightarrow YZ}{X \not\longrightarrow Y} \qquad (7) \quad \frac{Y \longrightarrow Z\ X \not\longrightarrow Z}{X \not\longrightarrow Y}$$

A (3) és (7) szabályok a (2) lehetséges megfordításai, a (4) és (5) pedig az (1) megfordításaiként tekinthetők. A (6) szabály az alábbi összevonási szabály megfordítása, mely az (1) és a (2) szabály alapján következik:

$$(8) \quad \frac{X \longrightarrow Y,\ X \longrightarrow Z}{X \longrightarrow YZ}$$

Egy axiómarendszerrel történő *levezethetőség*et általában a $\vdash$ szimbólum jelöli, a mi esetünkben például $\mathcal{F} \vdash \delta$ azt jelenti, hogy az $\mathcal{F}$ függőségi halmazból valamely (helyes és teljes) szabályrendszerrel levezethető a δ függőség. Mivel többféle axiomatizációt vezetünk be, jelezni fogjuk azt is, hogy mely rendszerben értendő a levezethetőség. Így az eredeti Armstrong-rendszerben történő levezethetőséget a $\vdash_A$, a kiterjesztett Armstrong-rendszerbelit pedig a $\vdash_{EA}$ fogja jelölni. Hasonlóképpen, a később bevezetett axiomatizációknál értelemszerű indexeket fogunk használni.

Tekintsünk egy példát a fenti szabályrendszerrel való következtetésre. Tegyük fel, hogy egy $\{$*Járat*, *Pilóta*, *Idő*$\}$ relációséma repülőjáratok és az azokat vezető pilóták összerendelésének sémáját írja le, most az idő feltüntetésével. A séma szintjén előírjuk, hogy egy pilóta egyidőben csak egy járatot vezethet: *Idő Pilóta* $\to$ *Járat*; továbbá azt is, hogy egy járatot egy időben csak egy pilóta vezet (a másodpilótát nem jegyezzük be ide): *Járat Idő* $\to$ *Pilóta*. Ha ezek után azt is előírnánk, hogy egy járat egy meghatározott időben megy, akkor a *Járat* $\to$ *Idő* függőséget kellene hozzávennünk a halmazhoz. Ez esetben a fenti szabályrendszerrel levezethetjük az implikált függőségeket. Első lépésben az (1) bővítési szabállyal megkapjuk a *Járat Pilóta* $\to$ *Idő* függőséget. Ezek után viszont nem tudunk úgy szabályt alkalmazni, hogy egy lépésben ne triviális vagy közel triviális eredményt kapjunk. Így pl. megkaphatjuk a bővítési szabály újbóli alkalmazásával a *Járat* $\to$ *Járat Idő* függőséget, mely nem ad új információt, viszont most már alkalmazhatjuk a (2)

[4]Ha negált függőségekről is szó van, akkor ezt a kiterjesztett rendszert nevezzük egyszerűen Armstrong-axiómarendszernek.

tranzitivitást, melynek eredménye releváns: *Járat* $\rightarrow$ *Pilóta*. Ha most megvizsgáljuk, hogy zárt-e már a halmaz, a válasz természetesen *nem*, hiszen még levezethető pl. a *Járat* $\rightarrow$ *Járat*, *Járat* $\rightarrow$ *Járat Pilóta*, *Járat* $\rightarrow$ *Idő Pilóta*, ... Bár ez utóbbiak ismét redundánsak (a korábban levezetettekhez képest nem hordoznak új információt), általános esetben nem könnyű látni, hogy rajtuk keresztül kaphatunk-e még releváns függőséget. Célszerű lenne például, ha a redundáns (és különösen a triviális) függőségekkel nem kellene foglalkoznunk, hiszen azok a gyakorlatban feleslegesen bonyolítják a függőségi halmazok vizsgálatát.

1.2.5. Függőségi rendszerek megadása

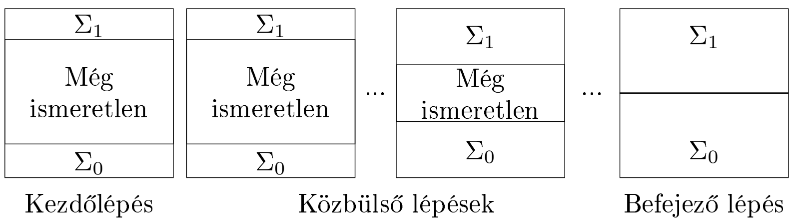
Tervezéskor – mint már említettük – fontos feladat az, hogy megállapítsuk, mely lehetséges függőségek igazak és melyek nem. Ez teszi lehetővé a teljes és egyértelmű specifikációt. Sajnos azonban nagyobb sémákra az ember nem képes áttekinteni a függőségi rendszereket, illetve a függőségek közti kapcsolatokat. Bár vannak elméleti módszerek a függőségi rendszerek kezelésére, egyelőre nem ismert olyan egyszerű reprezentáció, amely könnyen áttekinthetővé tenné a függőségek halmazait, kiemelné a hiányzó (nem specifikált érvényességű) függőségeket, illetve a függőségek közti logikai viszonyokat.

A függőségek érvényességének lépésenkénti megállapításához kétféle függőségi halmazt tarthatunk nyilván ([25, 28, 30]) – a korábban tárgyaltakkal összhangban:

A pozitív függőségek Σ_1 halmaza: Minden olyan funkcionális függőség, amelyet eleve igaznak tekintünk, illetve amelyek érvényessége következik a kezdetben megadott érvényes (pozitív) függőségekből (mint láttuk, negatívból pozitív nem következik).

A negatív függőségek Σ_0 halmaza: Minden olyan funkcionális függőség, amelyet eleve hamisnak tekintünk, illetve amelyek érvénytelensége következményként előáll a kezdetben megadott érvényes (pozitív) és érvénytelen (negatív) függőségekből kiindulva.

Ez a megközelítés az alábbi, az 1. ábrán is látható egyszerű módszerhez vezet, mely többféleképpen is finomítható; az ismert eljárások döntő többsége is ezen alapul. A módszer lényege, hogy minden lépésben egy még nem ismert érvényes-



1. ábra. Egy függőségi rendszer elkészítése lépésenként ségű lehetséges függőséget a modellezendő valósággal összevetve megvizsgálunk, és megállapítjuk annak érvényességét vagy érvénytelenségét, majd képezzük ennek és a korábban már eldöntött függőségeknek a logikai következményeit.

1. *Kezdő lépés: Rögzítjük a kezdeti függőségeket.*
2. *Iterált lépés: Ismételjük az alábbi lépéseket, amíg a Σ_0 vagy a Σ_1 halmaz változik:*
   - *Keresünk egy α lehetséges függőséget, mely nem eleme $\Sigma_0 \cup \Sigma_1$-nek.*
     - *Ha α érvényes (pozitív), akkor Σ_1-et bővítjük α-val.*
     - *Ha α nem érvényes (negatív), akkor Σ_0-t bővítjük α-val.*
   - *Képezzük a Σ_0 és a Σ_1 lezártját.*

1.3. A felmerülő problémák és az általunk készített reprezentáció

A funkcionális függőségek elmélete bár jól kidolgozott, a tervezési gyakorlatban gyakran csak közvetve alkalmazzák (pl. kulcsok, dekompozíciók, normálformák). Egyes esetekben viszont az egyértelmű és teljes specifikáció csak függőségi rendszer megadásával lehetséges. Ilyen például három vagy több egyedosztály kapcsolattípusának megadása, az egyed-kapcsolat modell ugyanis meglehetősen szűk lehetőségeket kínál többágú kapcsolatok jellegének deklarálására. A bináris kapcsolat jól ismert három típusa (egy-egy, egy-sok, sok-sok) megfogalmazható funkcionális függőségekkel. Emellett azonban a ternáris vagy nagyobb aritású kapcsolatok típusait nem elemzik kellőképpen, pedig rendelkezésre áll hozzá a függőségek elmélete (a gyakorlatban ritka, hogy nagy számú, pl. ötnél több egyedosztály egy kapcsolatban részt vegyen, de a kapcsolattípusok kisebb aritásokra sincsenek kellően kidolgozva). Ez esetenként többértelmű specifikációhoz vezethet, gyakrabban a kapcsolatok „binarizálásához” (a magasabb aritású kapcsolatok binárisakkal történő modellezése, közbülső egyedosztályok bevezetésével).

Ez többek között annak köszönhető, hogy a függőségeket hagyományosan egyenként kezeljük, a függőségi rendszereket pedig önálló függőségek halmazaiként. A közöttük lévő kapcsolatokat viszont kezelni kell; a logikai nyelv például lehetővé teszi az implikációt, de a formalizmus használata nehézkessé válik a gyakorlatban a vizuális áttekinthetőség hiánya és a redundanciák miatt (ld. az ismertetett implikációs példát a repülőjáratokkal). Az MFDBS'87 konferencián merült fel [29], hogy ehelyett inkább függőségek halmazaiban, a függőségi rendszer egészében kellene gondolkodni, és megfelelő specifikációs módszert adni a függőségi rendszerekhez, mely természetes, áttekinthető módon támogatja a halmazon belüli levezetést.

Ennek nyomán a funkcionális függőségi rendszereknek egy olyan diagramos reprezentációját adjuk, mely lehetővé teszi a hagyományos jelölésnél egyszerűbb kezelésüket, áttekintésüket. A fent említett algoritmusséma a mi diagramos rendszerünkben is alkalmazható, a grafikus reprezentáció kisebb sémákra az implikált függőségek levezetését közvetlenül a diagramon is lehetővé teszi. Később látni fogjuk, hogy a lezárás szisztematikusan elvégezhető. Ehhez a formalizmus egyszerűsítése szükséges; a hozzá kifejlesztett axiomatizáció és a levezetési algoritmus természetesen érvényes nagyobb sémákra is, bár ott a grafikus reprezentáció általánosítása bonyolulttá válik.

A lehetséges függőségek száma exponenciális az attribútumok számában [19]. Ezért az összes függőség érvényességének vagy érvénytelenségének specifikálása nem mindig valósítható meg, a tervezést korlátozzák kognitív képességeink (ezt jól példázza, hogy a grafikus reprezentáció nagyobb sémákra magasabb dimenziós terekben általánosítható egyszerűen). E probléma szorosan összefügg azzal az ismert kombinatorikai problémával, mely az MFDBS'87 konferencián merült fel ([29]), és csak részben megoldott: Egy n attribútumos sémán vajon hány különböző funkcionális függőségi rendszer (zárt halmaz) adható meg? Egy másik hasonló kérdés, hogy hányféle, funkcionális függőségekkel leírható n-ágú kapcsolat létezik (az egyed-kapcsolat diagram nyomán). Ez az előzőtől annyiban különbözik, hogy az attribútumok permutációjával egymásba vihető függőségi rendszerek azonosnak tekintendők. Kutatásunk során e problémával is foglalkoztunk, s legfeljebb 5 attribútum esetén a lehetséges függőségi rendszerek, ill. típusok számát meghatároztuk.

2. A funkcionális függőségi rendszerek szintaxisa

2.1. A függőségek alaphalmazai

Célunk egyelőre az, hogy elkerüljük a funkcionális függőségek hagyományos jelöléséből adódó redundanciákat, bevezetve egy egyszerűsített formalizmust, melyhez azután helyes és teljes következtetési szabályrendszert adunk. Az egyszerűsítés lényege, hogy eltekintünk a redundáns (és köztük a triviális) függőségektől, valamint csak kanonikus (szingleton) függőségekkel foglalkozunk.

Bevezetünk néhány további jelölést. A korábbiakkal összhangban a funcionális függőségek alaphalmazát, továbbá a nemredundáns, illetve kanonikus (szingleton) nemredundáns funkcionális függőségek alaphalmazát jelölje rendre $\mathbb{D}$, $\mathbb{D}^+$ és $\mathbb{D}_c^+$. Hasonlóan, $\mathbb{E}$, $\mathbb{E}^+$ és $\mathbb{E}_c^+$ jelöli az általános negált, a nemredundáns negált és a nemredundáns negált kanonikus (szingleton nemredundáns negált) függőségek alaphalmazát. Az Armstrong-rendszer alaphalmaza – mint láttuk – $\mathbb{D} \cup \mathbb{E}$, ezt fogjuk egyszerűsíteni $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$-ra. A függőségekre és halmazaikra vonatkozóan bevezetett fogalmakat (lezárt, zártság, teljesség, lehetséges kiterjesztés, logikai következmény) mindig az éppen adott univerzum fölött tekintjük, a levezethetőséget pedig jelölni fogjuk, hogy mely axiómarendszerben értendő.

Anélkül, hogy érdemi kifejezőerőt veszítenénk, rendszerünkből kihagyhatjuk a redundáns függőségeket (gondolva itt mind a pozitívakra; mind a negatívakra), azaz amelyek jobb és bal oldalán van közös attribútum, vagy a jobb oldaluk üres. Ezen kívül a nem kanonikus pozitív funkcionális függőségeket is elhagyhatjuk, hiszen azok helyettesíthetők több kanonikus függőséggel.[5] Sajnos azonban a bemutatott, kiterjesztett Armstrong-axiómarendszer nem felel meg ezeknek a megszorításoknak. Meg fogjuk mutatni, hogy egy vele ekvivalens rendszer készíthető (ld. később az ST és NST rendszert, illetve az 5.1. és 5.3. tételt) az egyszerűsített formalizmusra. Kiderül, hogy nincs szükség redundáns függőségek levezetésére ahhoz, hogy egy függőségi halmaz összes nemredundáns következményét levezessük. Így az első, közbülső egyszerűsítő lépés az eredeti univerzum $\mathbb{D}_c^+ \cup \mathbb{E}^+$-re való megszorítását jelenti, továbbá azt, hogy egy $\mathbb{D} \cup \mathbb{E}$ fölötti tetszőleges, véges $\mathcal{F}$ függőségi halmaz helyett a vele $\mathbb{D}_c^+ \cup \mathbb{E}^+$ fölött ekvivalens

$$\begin{aligned}\mathcal{F}' =&\{V \to C \mid \exists W : V \to W \in \mathcal{F},\, C \in W \setminus V\} \cup \\ &\cup \{V \not\to U \mid \exists W : V \not\to W \in \mathcal{F},\, U = W \setminus V\}\end{aligned}$$

függőségi halmazzal dolgozunk.

Az egyszerűsítés második lépése a nemszingleton negált függőségek (negált nemkanonikus függőségek) elhagyása. Bár ez a kifejezőerő valódi megszorítását jelenti, ennek a ténynek kevés gyakorlati jelentősége van. A nemszingleton negált függőségek ugyanis vagylagosságot fejeznek ki: pl. $X \not\to AB$ azt jelenti, hogy vagy $X \not\to A$, vagy $X \not\to B$ (vagy mindkettő) érvényes. Nyilvánvaló, hogy egy teljes függőségi halmazban az ilyen vagylagosságok redundánsak, hiszen minden lehetséges elemi, azaz kanonikus függőségre meg van határozva, hogy érvényes-e vagy sem – utóbbi esetben a negált szerepel a halmazban. Ha pedig az ilyen függőségek jobb oldala más függőségek miatt nem egyszerűsíthető le egyattribútumosra, azaz sem $X \to A$, sem $X \to B$ nem vezethető le, akkor csupán információhiányt vonhatunk le következtetésként: nem tudjuk eldönteni, hogy a két pozitív függőség közül egyáltalán igaz lehet-e valamelyik (csak azt tudjuk, hogy mindkettő egyszerre nem igaz).[6] Figyelembe véve, hogy sématervezéskor a cél egy teljes rendszerhez való eljutás, és a kezdeti halmazban általában nem szerepelnek ilyen nemszingleton negált függőségek, elfogadhatjuk, hogy elhagyásukkal a függőségek kifejezőerejét csekély módon csökkentsük. Ez a második egyszerűsítő lépés formálisan az alaphalmaz $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$-re való megszorítását jelenti, azaz az előzőekhez képest csak olyan függőségi halmazokkal foglalkozunk, melyekben nem szerepel $V \not\to W$, $|W| > 1$ alakú negált függőség.

A kiterjesztett Armstrong-axiómarendszer lehetővé teszi, hogy a $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ alaphalmazon kívül eső függőségeket is levezessünk még olyan halmazból indulva is ki, mely része ezen alaphalmaznak. Garantálnunk kell tehát, hogy ezek a közbülső eredmények ne legyenek szükségesek ahhoz, hogy $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$-beli következményeket vezessünk le. Látni fogjuk, hogy készíthető olyan szabályrendszer, mellyel minden szingleton következmény levezethető nemszingleton részeredmények nélkül, ha a kezdeti függőségi halmaz csak szingleton függőségeket tartalmaz. Ez lesz majd a PQRST axiómarendszer, melynek megvan az az előnyös tulajdonsága is, hogy

[5]Például az $X \to AB$ függőség helyettesíthető $X \to A$-val és $X \to B$-vel.

[6]Ha például a zárt világ feltételezést alkalmazzuk, akkor pedig az a konklúzió, hogy egyik sem igaz.

a szabályok alkalmazásának van egy kötött sorrendje, mely minden esetben teljes lezárást eredményez, azaz minden következmény levezethető úgy is, hogy a szabályokat csak egy előre megadott sorrendben alkalmazzuk.

2.2. A Dimenzió fogalma

Egy $X \to A \in \mathbb{D}_c^+$ funkcionális függőségre definiáljuk annak *dimenzióját*, melyet $[X \to A]$-val jelölünk:

$$[X \to A] \stackrel{\text{def}}{=} |X|$$

(a dimenzió negált függőségekre hasonlóan definiálható). Egy A attribútumra is értelmezzük a dimenzió fogalmát, adott $\mathcal{F} \subset \mathbb{D}_c^+$ függőségi halmaz mellett:

$$[A]_{\mathcal{F}} \stackrel{\text{def}}{=} \min_{X \to A \in \mathcal{F}^+} |X|,$$

ahol $\mathcal{F}^+$ az $\mathcal{F}$ lezártja a $\mathbb{D}_c^+$ fölött, azaz

$$\mathcal{F}^+ = \{\delta \in \mathbb{D}_c^+ \mid \mathcal{F} \models \delta\}.$$

A definíciót kibővítjük arra az esetre is, ha $\mathcal{F}^+$-ban nincs $X \to A$ alakú függőség, ekkor

$$[A]_{\mathcal{F}} \stackrel{\text{def}}{=} \infty.^7$$

A dimenzió fogalma megkönnyíti a függőségi halmazok kezelését, csoportosítását, szoros kapcsolata van a grafikus reprezentációval, és a következtetési szabályok alkalmazási sorrendjénél is van szerepe.

3. A függőségi halmazok száma

Jelölje n a séma attribútumainak számát, az $\mathcal{SD}_n$-nel jelölt halmaz pedig tartalmazza az összes olyan érvényes (zárt, ellentmondásmentes) funkcionális függőségi halmazt, melyben a függőségek bal oldala nem üres.[8] Definiálunk egy τ ekvivalenciarelációt ezen a halmazon oly módon, hogy két függőségi halmazt akkor tekintünk ekvivalensnek, ha létezik az attribútumok egy olyan permutációja, mely az egyik halmazt a másikba viszi át. Az ekvivalenciaosztályok (a függőségi rendszerek *típusai* v. *esetei*) halmazát $\mathcal{SD}_n/\tau$ jelöli, ennek a méretére koncentrálunk. Ha

[7] Ezt az esetet definiálhatnánk úgy is, hogy $[A] = n$ (ahol n a séma attribútumainak száma). A ∞ jelölés viszont hangsúlyozza, hogy A független a többi attribútumtól, továbbá maga a dimenzió definíciója függetlenné válik n-től.

[8] Ennek a megkötésnek a sématervezésben van szerepe. Például egy egyed-kapcsolat modellben a többágú kapcsolatoknak megfeleltethetünk függőségi halmazokat (ha a multiplicitások megfelelő alakúak, és a kapcsolatokban lehetnek belső multiplicitási megszorítások is), s ilyenkor nem fordulnak elő konstans attribútumokat megadó függőségek.

megengedjük azokat a függőségeket is, amelyek konstansnak deklarálnak egy-egy attribútumot (azaz üres halmaz áll bal oldalukon), akkor egy bővebb halmazt kapunk, melyet $\mathcal{SD}_n^0$ jelöl, ennek különböző eseteit pedig $\mathcal{SD}_n^0/\tau$. Könnyen látható, hogy

$$\left|\mathcal{SD}_{n+1}^0/\tau\right| = \left|\mathcal{SD}_{n+1}/\tau\right| + \left|\mathcal{SD}_n^0/\tau\right|$$

teljesül minden $n \in \mathbb{N}^+$-ra.

Ezekkel a jelölésekkel táblázatba foglaltuk a lehetséges függőségi halmazok és esetek számát, öt vagy annál kevesebb attribútum esetén. Az eredményeket egy PROLOG program segítségével kaptuk [22], mely az (S) és (T) szabályokra, s azok alkalmazási sorrendjére épül (ld. 5.1. tétel). Magukat a lehetséges eseteket is tartalmazza [22] – amennyire a terjedelem engedi.

Egyelőre nyitott kérdés maradt az ötnél több attribútumot tartalmazó függőségi halmazok lehetséges száma.

n	$\|\mathcal{SD}_n\|$	$\|\mathcal{SD}_n/\tau\|$	$\|\mathcal{SD}_n^0\|$	$\|\mathcal{SD}_n^0/\tau\|$
1	1	*1*	2	2
2	4	*3*	7	5
3	45	*14*	61	19
4	2 271	*165*	2 480	184
5	1 373 701	*14 480*	1 385 552	14 664

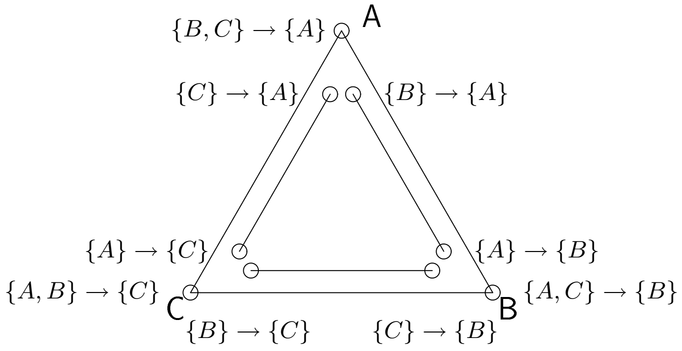
1. táblázat. n attribútumos sémák lehetséges zárt függőségi halmazainak száma.

4. Funkcionális függőségi rendszerek ábrázolása diagramokon

A függőségi rendszerek ábrázolására már fejlesztettek ki módszereket. Ilyen például [3] vagy [33] gráfelméleti megközelítése. Ezek és a hasonló módszerek nem terjedtek el a gyakorlatban, főként azért, mert bonyolultabb függőségi rendszerek esetén kevésbé áttekinthetők. A háromattribútumos esetre [11] javasol egy lehetséges diagramot, amely azonban redundáns, és nem általánosítható nagyobb sémákra. A mi jelölésünk valamelyest egyszerűbb; könnyen látható rajta, mely függőségek igazak és melyek nem, ezen kívül következtetésre is alkalmas.

4.1. Háromszöges reprezentáció

A 2. ábrán láthatjuk a diagram sémáját három attribútum esetén. Ebben az esetben alapvetően kétféle függőséget különböztetünk meg. A függőségeket a háromszög csúcspontjai és a mellé rajzolt szakaszok (külön élek) végpontjai reprezentálják:



2. ábra. Háromszöges diagram háromattribútumos sémák funkcionális függőségi rendszereinek ábrázolásához. A különálló szakaszok a háromszögön belülre vagy kívülre egyaránt rajzolhatók

- *Egydimenziós függőségek (a bal oldal egy attribútumot tartalmaz):*
  Az $\{A\} \to \{B\}$ függőség például ilyen, a diagramon a háromszög AB éle mentén lévő szakasz (mint egydimenziós alakzat) B felőli végpontja reprezentálja.
- *Kétdimenziós függőségek (két attribútum a bal oldalon):*
  Például az $\{A,B\} \to \{C\}$ függőség tartozik ebbe a csoportba, ennek a háromszög (kétdimenziós alakzat) C felőli csúcsa felel meg a diagramon.

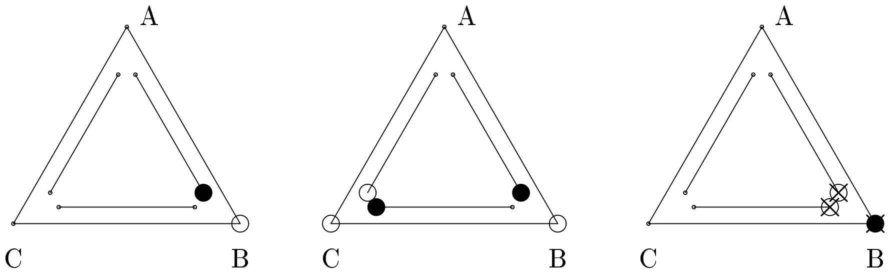
Ha egy függőség érvényes, azt a neki megfelelő csúcspontba rajzolt körrel ábrázoljuk. Áthúzott körrel jelöjük, ha tudjuk, hogy az adott függőség nem (vagy nem mindig) teljesül, azaz a neki megfelelő negált függőség eleme a függőségi rendszernek. Kis körrel jelöljük, ha egy függőség érvényessége nem állapítható meg a kiindulási halmaz alapján. Ha negatív függőségeket nem kezelünk és lezártuk a pozitív függőségek halmazát, a fennmaradó csúcsokra szintén kis körök kerülhetnek.

Az implikációt is figyelembe véve összességében az alábbi jelöléseket alkalmazhatjuk a diagramon:

- A kezdetben megadott függőségeket *tele körökkel*,
- az implikált függőségeket *üres körökkel*,
- a kezdeti negált függőségeket *áthúzott tele körökkel* (vagy kis körökkel) és
- az implikált negált függőségeket *áthúzott üres körökkel* (vagy kis körökkel) jelöljük.

Néhány példát láthatunk a 3. ábrán. A bal oldali az $\{A\} \to \{B\}$ függőséget ábrázolja az $\{A,C\} \to \{B\}$ implikált függőséggel. A középső az

$$\{A\} \to \{B\}, \quad \{B\} \to \{C\}$$



3. ábra. Példák a háromszöges reprezentációra

függőségek és következményeik, a jobb oldali háromszögön pedig az $\{A,C\} \not\rightarrow \{B\}$ negált függőség és $\{A\} \not\rightarrow \{B\}$, $\{C\} \not\rightarrow \{B\}$ következményei láthatók.

Az összes lehetséges eset diagramját [22] tartalmazza.

4.2. Általánosítás több attribútumra

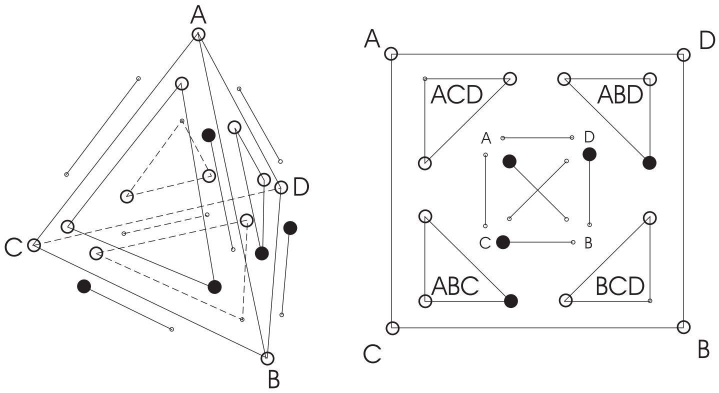
A fenti háromszöges reprezentáció általánosítható több attribútumot tartalmazó sémákra is. Két irány képzelhető el.

A dimenziófogalmat követve magasabb dimenziós terekben folytatható az építkezés a háromszöges alapon: pl. négy attribútum esetén *tetraéderes* reprezentáció adható a háromdimenziós térben. Ilyenkor a diagram egy tetraéderből, négy háromszögből (beágyazott háromattribútumos részek) és hat szakaszból (élből) áll (beágyazott bináris részek).

A másik lehetséges irány az, hogy síkban maradunk: a tetraéder négyzetre történő kicserélésével és a többi elem megfelelő elrendezésével kapjuk a *négyzetes* reprezentációt. A párhuzamosság és az irányok figyelembe vételével könnyen végiggondolható, mely síkidom mely csúcsa minek felel meg a tetraéderes reprezentációban.

A négyattribútumos esetre láthatunk példát a 4. ábrán (az implikációt később taglaljuk).

A konstrukció általánosítható 5 attribútumra is. A kombinatorikus robbanás miatt azonban az attribútumok számának növelésével a reprezentáció bonyolulttá válik, több geometriai alakzatra van szükség. n attribútum esetén a viszonylag egyszerű struktúrájú, szimmetrikus reprezentáció az $n-1$ dimenziós térben létezik – ennek átvitele két dimenzióba a szimmetriák megtartásával nem lehetséges. A nagyobb sémák bonyolultsága az attribútumok csoportosításával, dekomponálásával kezelhető, vagy az egész séma helyett egyes transzlációk ([17], [18]) ábrázolásával. A [23] cikkben bemutatunk egy példát, mely az attribútumok csoportosítására épül. Az ilyen lehetőségek részletesebb vizsgálata további kutatómunkát igényel, összefüggésben a kis sémák már ismert lehetséges függőségi rendszereinek kategorizálásával, jellemzésével.



4. ábra. A $\{B \to C, B \to D, B \to A, AD \to B, AC \to B\}$ által generált rendszer tetraéderes és négyzetes reprezentációja.

Ha üres bal oldalú függőségek is megengedettek (azaz konstans attribútumok is megadhatók), akkor a diagramokat attribútumonként egy-egy további ponttal kell bővíteni, mely nulla dimenziós elemként éppen ezen függőségek reprezentációját teszi lehetővé. A szabályok grafikus mintáit is bővíteni kell az $Y = \emptyset$ esettel, mely könnyen elvégezhető.

5. Axiómarendszerek

5.1. Az ST és a PQRST axiómarendszer

Az alábbi levezetési szabályokban és axiómában Y egy attribútumhalmazt jelöl (mely akár üres is lehet), A, B és C pedig olyan, páronként különböző attribútumokat, melyek nem szerepelnek Y-ban.

$$(S)\quad \frac{Y \to B}{YC \to B} \qquad (T)\quad \frac{Y \to A,\, YA \to B}{Y \to B} \qquad (P)\quad \frac{YC \not\to B}{Y \not\to B}$$

$$(Q)\quad \frac{Y \to A,\, Y \not\to B}{YA \not\to B} \qquad (R)\quad \frac{YA \to B,\, Y \not\to B}{Y \not\to A} \qquad (\square)\quad \neg(Y \to B,\, Y \not\to B)$$

Ha a dimenzionalitásra gondolunk, láthatjuk, hogy a szabályok Y méretének megfelelő dimenziójú függőségeken dolgoznak, beleértve esetleg eggyel nagyobb dimenziójúakat is. Ennek a ténynek fontos szerepe van a függőségek grafikus reprezentációiban való következtetésben, valamint a szabályok alkalmazási sorrendjének finomhangolásában. A szabályok Y különböző méretei mellett különböző dimenziójú eseteket jelölnek, melyek a grafikus reprezentációban könnyen felismerhetők és alkalmazhatók. Az (S) szabályt kiterjesztési szabálynak, a (T)-t pedig redukciós vagy forgatási szabálynak nevezhetjük. Pozitív (nem negált) függőségekre belátható, hogy ha először csak az (S)-t alkalmazzuk ameddig lehet, majd utána ugyanígy a (T)-t, megkapjuk az összes következményt (ld. az 5.1. tételt lejjebb). Ezt finomíthatjuk azzal, hogy (S)-t először alacsonyabb dimenzióban, majd fokozatosan magasabb dimenzióban alkalmazzuk, a (T) alkalmazását pedig a lehető legmagasabb dimenzióban kezdjük, s haladunk a legalacsonyabb felé. Ezt használtuk föl arra, hogy legeneráljuk, és összeszámoljuk a különböző eseteket (a függőségi halmazok típusait) négy ill. öt attribútum esetén (ld. 4. fejezet).

Ha negált függőségek is vannak, ehhez hozzá kell vegyük a (P), (Q) és (R) szabályokat is, mint az (S) és (T) negáltjait. A ($\square$) szimbolikusan azt fogalmazza meg, hogy „$\not\to$" épp a „$\to$" negáltja, azaz egy ellentmondásos függőségi halmazból levezethető a $\neg(\square)$. Mint fent, itt is kapunk a szabályok alkalmazására egy teljes módszert; melynek során minden egyes szabályt annyiszor alkalmazunk, ahányszor csak lehet, s utána térünk rá a következő szabályra. A sorrend: (S), (T), (R), (P), (Q); de (P) és (Q) felcserélhető.

Megjegyezzük, hogy csak szingleton, nemredundáns függőségekkel foglalkozunk. A szabályok alkalmazásával nem is léphetünk ki ebből a körből. Ez igaz mind a pozitív, mind a negált függőségekre.

A bemutatott szabályokból az alábbi két axiómarendszert definiáljuk; az elsőt funkcionális függőségekre, a másodikat arra az esetre, amikor pozitív és negált függőségek is előfordulnak:

- *ST axiómarendszer:* a $\mathbb{D}_c^+$ alaphalmazon, az (S) és (T) szabályokkal (az Armstrong-rendszerbeli (0)-nak megfelelő axiómaséma nincs),
- *PQRST axiómarendszer:* a $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$alaphalmazon, az összes szabállyal és a ($\square$) szimbolikus axiómával, mely az ellentmondásosság jelzésére szolgál.

Ezek a rendszerek helyesek és teljesek alaphalmazukon. Ezt és a szabályok említett alkalmazási sorrendjét fogalmazza meg az alábbi két tétel.

5.1. Tétel. *Legyen $\mathcal{F}$ a $\mathbb{D}_c^+$ egy véges részhalmaza, továbbá $\delta \in \mathbb{D}_c^+$.*

(i) *Az ST axiómarendszer helyes és teljes a $\mathbb{D}_c^+$ alaphalmaz fölött, azaz $\mathcal{F} \vdash_{ST} \delta \iff \mathcal{F} \models \delta$.*

(ii) *Ha $\mathcal{F} \vdash_{ST} \mathcal{G}$ egy véges $\mathcal{G} \subset \mathbb{D}_c^+$ függőségi halmazra, akkor $\mathcal{G}$ összes eleme levezethető $\mathcal{F}$-ből az (S) és (T) szabályok alkalmazásával úgy, hogy (S) egyetlen alkalmazását sem előzi meg (T) alkalmazása.*

5.2. Tétel. *Legyen $\mathcal{F}$ a $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ egy véges részhalmaza, továbbá $\delta \in \mathbb{D}_c^+ \cup \mathbb{E}_c^+$.*

(i) *A PQRST axiómarendszer a ($\square$) nélkül helyes $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ fölött, és teljes minden ellentmondásmentes $\mathcal{F}$-re, azaz $\mathcal{F} \vdash_{PQRST} \delta \iff \mathcal{F} \models \delta$, ha*

$\mathcal{F}$ ellentmondásmentes. Ezen kívül, $\neg(\Box)$ levezethető akkor és csak akkor, ha $\mathcal{F}$ ellentmondásos.[9]

(ii) *Ha $\mathcal{F} \vdash_{PQRST} \mathcal{G}$ egy véges $\mathcal{G} \subset \mathbb{D}_c^+ \cup \mathbb{E}_c^+$ függőségi halmazra, akkor $\mathcal{G}$ összes eleme levezethető $\mathcal{F}$-ből az (S), (T), (R), (P), (Q) szabályok alkalmazásával oly módon, hogy (R) egyetlen alkalmazását sem előzi meg (P) és (Q) alkalmazása, (T) alkalmazását nem előzi meg (R) alkalmazása, valamint (S) alkalmazását nem előzi meg (T) alkalmazása. Ezen kívül, ha $|\mathcal{G}| = 1$, akkor (R)-et elegendő egyszer alkalmazni.*

A tételek bizonyítása az Armstrong-axiómarendszer $\mathbb{D}$ fölötti és a kiterjesztett változat $\mathbb{D} \cup \mathbb{E}$ fölötti helyességén és teljességén múlik. A következőkben kiterjesztjük az alaphalmazt a nemszingleton negált függőségekkel (így $\mathbb{D}_c^+ \cup \mathbb{E}^+$-t kapunk), s bevezetünk egy másik axiómarendszert, az NST-t. Belátható, hogy ez ekvivalens a PQRST-vel az eredeti, szűkített alaphalmazra ($\mathbb{D}_c^+ \cup \mathbb{E}_c^+$) nézve. Ezek után bevezetjük az U és UE szabályrendszereket, melyek rendre ekvivalensek az ST-vel és az NST-vel. Végül az U és UE rendszereket hasonlítjuk össze az eredeti Armstrong- és a kiterjesztett Armstrong-axiómarendszerekkel, így válik teljessé a bizonyítás menete.

5.2. Az NST axiómarendszer mint a PQRST bővítése

Az alaphalmazt kibővítjük $\mathbb{D}_c^+ \cup \mathbb{E}^+$-re.

Az alábbi szabályrendszerben Y, Z és W páronként diszjunkt attribútumhalmazokat jelölnek (melyek üresek is lehetnek azzal a megkötéssel, hogy a negált függőségek jobb oldalai nem lehetnek teljesen üresek), továbbá A, B és C különböző attribútumok, melyek nem fordulnak elő a velük egy szabályban szereplő halmazokban.

$$(S) \quad \frac{Y \to B}{YC \to B} \qquad\qquad (T) \quad \frac{Y \to A,\ YA \to B}{Y \to B}$$

$$(NS) \quad \frac{YC \not\to Z}{Y \not\to Z} \qquad\qquad (NT1) \quad \frac{Y \to A,\ Y \not\to Z}{YA \not\to Z}$$

[9] Az ellentmondásosság azt jelenti, hogy két ellentétes függőség (pl. $X \to W$, $X \not\to W$, lehetnek akár redundánsak is) levezethető az Armstrong-axiómarendszerben. Elegendő megmutatnunk, hogy ilyenkor $\neg(\Box)$ levezethető a PQRST rendszerben. Egyes ellentmondásos esetekben ugyanis az Armstrong-axiómarendszerben levezethető néhány nemredundáns, szingleton függőség, melyek nem vezethetőek le a PQRST-ben, az alaphalmaz megszorítottsága miatt (pl. a (6)-os szabállyal levezetjük $X \not\to \emptyset$-t, és utána bármely B-re a (4)-es szabállyal levezethető $X \not\to B$). Bár ezek az esetek nem relevánsak, szigorú formális értelemben vett teljesség nem állítható.

$$(NT2)\quad \frac{YZW \to B,\, Y \not\twoheadrightarrow ZB}{Y \not\twoheadrightarrow ZW} \qquad\qquad (NT2S)\quad \frac{Y \to B,\, Y \not\twoheadrightarrow ZB}{Y \not\twoheadrightarrow Z}$$

$$(\square)\quad \neg(Y \to B,\, Y \not\twoheadrightarrow B) \qquad\qquad (*)\quad \frac{Y \not\twoheadrightarrow Z}{Y \not\twoheadrightarrow ZC}$$

Ezekkel a szabályokkal akkor is működik a levezetés, ha nemszingleton negált függőségeket is figyelembe kell vennünk. Pozitív, funkcionális függőségek levezetésére továbbra is az (S) és a (T) szabály használható, de negáltjukként most az (NS), az (NT1), az (NT2) és a $(*)$ jelenik meg. Bár az (NT2S) levezethető (S)-ből és (NT2)-ből a $W = \emptyset$ választással (ld. az 5.8. lemmát később), az eset gyakran előfordul, így érdemes külön szabályként szerepeltetni.

A gyakorlatban a $(*)$ szabály nem szükséges, mert minden lényeges információt levezethetünk nélküle (ld. az 5.3. tétel 2. részét). Ez azt jelenti, hogy sosem kell bővítenünk egy negált függőség jobb oldalát, ami különösen kedvező, hiszen általában épp az a célunk, hogy ezeket a jobb oldalakat redukáljuk. Így egyedül az (NT2) szabály esete az, ami a jobb oldal méretét növelheti, de ez azzal jár, hogy közben legalább egy attribútumot el is távolítunk belőle ($B \notin W$).

Definiáljuk az *NST axiómarendszert* a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ alaphalmazon mint a fenti szabályok együttesét (az (NT2S) opcionális), beleértve a $(\square)$ ellentmondást jelző szimbolikus axiómát is. A rendszer helyes és teljes a nemredundáns függőségekre vonatkozóan, ha a pozitív függőségek kanonikusak. Jelöljük továbbá *NST'*-vel azt az axiómarendszert, melyet NST-ből $(*)$ elhagyásával kapunk.

5.3. TÉTEL. *Legyen $\mathcal{F}$ a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ egy véges részhalmaza, továbbá $\delta \in \mathbb{D}_c^+ \cup \mathbb{E}^+$.*

(i) Az NST rendszer $(\square)$ nélkül helyes a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ alaphalmazon és teljes minden ellentmondásmentes $\mathcal{F}$-re, azaz $\mathcal{F} \vdash_{NST} \delta \iff \mathcal{F} \models \delta$, ha $\mathcal{F}$ ellentmondásmentes. Ezen kívül, $\neg(\square)$ levezethető akkor és csak akkor, ha $\mathcal{F}$ ellentmondásos.

(ii) A $(*)$ szabály nem szükséges ahhoz, hogy az NST rendszerben minden releváns információt levezessünk, azaz az NST' rendszer is helyes a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ alaphalmazon, teljes $\mathbb{D}_c^+$-n; és ha $\mathcal{F}$ ellentmondásmentes, továbbá $\delta \in \mathbb{E}^+$, $\delta = X \not\twoheadrightarrow Y$, akkor

$$\mathcal{F} \models \delta \Longrightarrow \exists Z \subseteq Y,\, Y \neq \emptyset : \mathcal{F} \vdash_{NST'} X \not\twoheadrightarrow Z.^{10}$$

Ha $\mathcal{F}$ ellentmondásos, akkor $\mathcal{F} \vdash_{NST'} \neg(\square)$.

Az 5.2. tétel bizonyítása az 5.3. tételen és az alábbi lemmákon nyugszik, melyek a PQRST és az NST $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ alaphalmaz fölötti ekvivalenciáját állítják.

[10] Vegyük észre, hogy $X \not\twoheadrightarrow Z$ legalább annyi információt hordoz, mint δ.

5.1. LEMMA. *Ha az NST helyes rendszer, akkor PQRST is helyes, azaz* $\vdash_{NST}\{(P),(Q),(R)\}$.

5.2. LEMMA. *Vegyünk egy függőségi halmazt, mely nemredundáns, szingleton (pozitív és negatív) függőségeket tartalmazhat. Ekkor minden olyan nemredundáns, szingleton függőség, mely NST'-ben levezethető* $\mathbb{D}_c^+ \cup \mathbb{E}^+$ *fölött, szintén levezethető a PQRST rendszerben* $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ *fölött, ha a kezdeti függőségi halmaz ellentmondásmentes. Ha ellentmondásos a halmaz, akkor pedig a PQRST rendszerben* $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ *fölött levezethető a* $\neg(\Box)$. *Pontosabban fogalmazva:*

- *Legyen* $\mathcal{F}$ *a* $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ *egy véges részhalmaza,* X *egy attribútumhalmaz és* A *egy attribútum, melyre* $A \notin X$. *Ha* $\mathcal{F}$ *ellentmondásmentes és* $\mathcal{F} \vdash_{NST'} X \not\rightarrow A$, *akkor* $\mathcal{F} \vdash_{PQRST} X \not\rightarrow A$. *Ha* $\mathcal{F}$ *ellentmondásos, akkor* $\mathcal{F} \vdash_{PQRST} \neg(\Box)$ *teljesül.*
- *Továbbá, ha* $\mathcal{F} \vdash_{PQRST} X \not\rightarrow A$, *akkor a levezetés elvégezhető úgy, hogy az (R) szabályt legföljebb egyszer kell alkalmazni.*

5.3. Az U és az UE axiómarendszerek és kapcsolatuk az ST, az NST és a kiterjesztett Armstrong-axiómarendszerekkel

Tekintsük az alábbi szabályokat. Csakúgy mint előbb, X, Y, Z és W páronként diszjunkt attribútumhalmazokat jelölnek azokra a szabályokra nézve, melyekben együtt előfordulnak (lehetnek üresek is, de a negált függőségek jobb oldala nem válhat üressé), továbbá az A_i-k ($i \in [1..k]$, $k \in \mathbb{N}_0$) különböző attribútumokat jelölnek, melyek nem elemei a velük egy szabályban szereplő halmazoknak.

$$(U) \quad \frac{XY \rightarrow A_1, \cdots, XY \rightarrow A_k,\ YA_1 \cdots A_k \rightarrow B}{XY \rightarrow B}$$

$$(E1) \quad \frac{XY \rightarrow A_1, \cdots, XY \rightarrow A_k,\ XY \not\rightarrow ZA_1 \cdots A_l\ (0 \le l \le k)}{YA_1 \cdots A_k \not\rightarrow Z}$$

$$(E2) \quad \frac{X \rightarrow A_1, \cdots, X \rightarrow A_k,\ X \not\rightarrow ZA_1 \cdots A_k}{X \not\rightarrow Z}$$

$$(E3) \quad \frac{XZW \rightarrow A_1, \cdots, XZW \rightarrow A_k,\ XY \not\rightarrow ZA_1 \cdots A_k}{XY \not\rightarrow ZW}$$

$$(E\Box) \quad \neg(X \rightarrow A_1, \cdots, X \rightarrow A_k,\ X \not\rightarrow A_1 \cdots A_k)$$

Az $\{(U)\}$ egyetlen szabályból álló, $\mathbb{D}_c^+$ fölött értelmezett rendszer *U axiómarendszernek* nevezzük. Kibővítve az (E1), (E2), (E3) szabályokkal és az (E□)-vel, megkapjuk az *UE axiómarendszert* a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ alaphalmazon.[11]

Megjegyzendő, hogy k (és l) lehet 0 is, az (E2) kivételével (ott ez az eset irreleváns). Megmutatható, hogy minden releváns információ levezethető anélkül, hogy (E3)-at $k = 0$-val alkalmazni kellene (ez az eset felel meg a $(*)$ szabálynak az

[11] Szoros értelemben ez nem véges axiomatizáció (de rekurzív), hiszen a felírt szabályok valójában sémák, melyekből minden k-ra kapunk egy-egy konkrét szabályt.

NST rendszerben). Ezt az állítást és az UE rendszernek a kiterjesztett Armstrong-axiómarendszerrel a $\mathbb{D}_c^+ \cup \mathbb{E}^+$ fölötti ekvivalenciáját mondják ki az alábbi lemmák. $\vdash_A$ jelöli az Armstrong-rendszerbeli bizonyíthatóságot a funkcionális függőségekre (értve ez alatt a (0) axiómát és az (1)-es és (2)-es szabályokat, $\mathbb{D}$ fölött), $\vdash_{EA}$ pedig a kiterjesztett Armstrong-rendszerbeli bizonyíthatóságot ($\mathbb{D} \cup \mathbb{E}$ fölött, az összes szabályt beleértve). E két rendszerről ismert, hogy helyesek és teljesek.

5.3. LEMMA. *(U) helyes, azaz* $\vdash_A (U)$.

5.4. LEMMA. *Az U szabályrendszer teljes a funkcionális függőségekre nézve a redundáns függőségek kivételével, továbbá azzal a megkötéssel, hogy azokat a függőségeket, melyek jobb oldala több attribútumból áll, kanonikus függőségek reprezentálják. Pontosabban: legyen* $\mathcal{F}$ *a* $\mathbb{D}$ *egy véges részhalmaza. Ekkor minden* X *és* Z *attribútumhalmazra*

$$\mathcal{F} \vdash_A X \to Z \Longrightarrow \forall B \in Z \setminus X : \mathcal{F}' \vdash_U X \to B,$$

ahol

$$\mathcal{F}' = \{V \to C \mid \exists W : V \to W \in \mathcal{F}, C \in W \setminus V\}.$$

5.5. LEMMA. *Az (E1), (E2), (E3) szabályok és (E□) helyesek, azaz* $\vdash_{EA}\{$*(E1), (E2), (E3)*$\}$ *és ha* $\mathcal{F} \vdash_{UE} \neg$*(E□) egy véges* $\mathcal{F} \subset \mathbb{D}_c^+ \cup \mathbb{E}^+$ *halmazra, akkor* $\mathcal{F}$ *ellentmondásos.*

5.6. LEMMA. *AZ UE rendszer teljes a negált funkcionális függőségekre nézve, a redundáns függőségek kivételével, ha a kiindulási függőséghalmaz ellentmondásmentes. Azokat a pozitív függőségeket, melyek jobb oldala több attribútumból áll, kanonikus függőségek reprezentálják (az 5.4. lemmához hasonlóan), továbbá a nem önellentmondásos redundáns negált függőségeket nemredundánsakkal helyettesítjük. Pontosabban:*

- *Legyen* $\mathcal{F}$ *a* $\mathbb{D} \cup \mathbb{E}$ *egy véges, ellentmondásmentes részhalmaza, és legyenek* X *és* Z *olyan attribútumhalmazok, melyekre* $Z \setminus X \neq \emptyset$ *teljesül. Így ha* $\mathcal{F} \vdash_{EA} X \nrightarrow Z$*, akkor* $\mathcal{F}' \vdash_{UE} X \nrightarrow Z \setminus X$*, ahol*

$$\begin{aligned}\mathcal{F}' =& \{V \to C \mid \exists W : V \to W \in \mathcal{F}, C \in W \setminus V\} \cup \\ & \{V \nrightarrow U \mid \exists W : V \nrightarrow W \in \mathcal{F}, U = W \setminus V\}.\end{aligned}$$

- *Továbbá, ha* $\mathcal{F}' \vdash_{UE} X \nrightarrow V$ *igaz egy nemredundáns negált függőségre (és* $\mathcal{F}'$ *nem ellentmondásos), akkor* $\exists U \subseteq V$, $U \neq \emptyset$ *úgy, hogy* $X \nrightarrow U$ *levezethető az UE rendszerben anélkül, hogy alkalmaznánk az (E3) szabályt* $k = 0$*-val.*
- *Ha pedig* $\mathcal{F}$ *ellentmondásos, és a hozzá tartozó* $\mathcal{F}'$*-re teljesül* $\mathcal{F}' \subset \mathbb{D}_c^+ \cup \mathbb{E}^+$ *(azaz* $\mathcal{F}$ *nem tartalmaz redundáns negált függőségeket, és így* $\mathcal{F}'$ *nem tartalmaz* $X \nrightarrow \emptyset$ *alakú függőségeket), akkor* $\mathcal{F}' \vdash_{UE} \neg$*(E□) igaz, és a levezetés elvégezhető anélkül, hogy (E3)-at* $k = 0$*-val alkalmaznánk.*

Az ST és az U, valamint az NST és az UE ekvivalenciájáról szónak az alábbi lemmák:

5.7. LEMMA. *Az ST és az U axiómarendszerek ekvivalensek a* $\mathbb{D}_c^+$ *fölött, azaz* $\vdash_U\{(S),\ (T)\}$ *és* $\vdash_{ST}(U)$.

5.8. LEMMA. *Az NST és az UE axiómarendszerek ekvivalensek a* $\mathbb{D}_c^+ \cup \mathbb{E}^+$ *fölött, azaz az 5.7. lemma állításain kívül az alábbiak igazak:*

- $\vdash_{UE}\{(NS),\ (NT1),\ (NT2),\ (*)\}$,
- $\{(S),\ (NT2)\}\vdash_{NST}(NT2S)$, $\neg(E\Box)\vdash_{NST}\neg(\Box)$,
- $\vdash_{NST}\{(E1),\ (E2),\ (E3)\}$, $\neg(\Box)\vdash_{UE}\neg(E\Box)$.
- *Ezen kívül a* $(*)$ *szabály csak az (E3) levezetéséhez szükséges* $k = 0$*-ra és fordítva (sem a többi szabály levezetéséhez, sem a (E*$\Box$*) levezetéséhez (*$\Box$*)-ból vagy fordítva).*

5.4. A bizonyítások

Az 5.1. tétel első részének bizonyítása az Armstrong-axiómák helyességén és teljességén, továbbá a 5.3., a 5.4. és a 5.7. lemmán nyugszik. A második részhez meg fogjuk mutatni, hogy egy (T)-t követő (S)-alkalmazás felírható mint egy (S)-et követő két (T)-alkalmazás.

A 5.2. tétel első része triviálisan következik a 5.3. tételből és az 5.1. és 5.2. lemmából. A második rész bizonyításához, hasonlóan az 5.1. tételhez, meg fogjuk mutatni, hogy bármely, a szabályokat nem a megadott sorrendben alkalmazott bizonyítás átalakítható egy olyan, vele ekvivalens bizonyítássá, melyben a szabályok követik a sorrendet (és ha a célhalmaz egyelemű, akkor (R)-et legföljebb egyszer kell alkalmazni).

Az 5.3. tétel az Armstrong-axiómák helyességén és teljességén nyugszik, valamint az 5.5., az 5.6. és a 5.8. lemmán.

A lemmák bizonyításai közül az 5.2., az 5.4. és az 5.6. érdekesebb, ezekben egy adott, tetszőleges levezetéssel párhuzamosan konstruálható meg a másik axiómarendszerben történő, megfelelő levezetés. Itt ugyanis a formális rendszer szintaxisában van különbség: az (U) rendszer például $\mathbb{D}_c^+$ fölötti, míg az eredeti Armstrong-axiomatizáció (a (0) axióma az (1) és (2) szabályokkal) $\mathbb{D}$ fölötti.

Először az 5.3. – 5.8. lemmákat bizonyítjuk, melyeket az 5.1. és az 5.2. követ. A tételek bizonyításait a lemmák után ismertetjük. A sorrend alapja a lemmák és a tételek logikai egymásra épülése.

Az 5.3. lemma bizonyítása. A $k = 0$ eset éppen az (1) szabálynak felel meg, ha $V = \emptyset$. Ellenkező esetben tetszőleges $k > 0$-ra tegyük fel, hogy

$$XY \to A_1, \ldots, XY \to A_k \text{ és } YA_1 \ldots A_k \to B$$

igaz (eleme a függőségi rendszernek). A (0) axióma szerint $XY \to Y$. A (8) szabály többszöri alkalmazásával $XY \to YA_1 \cdots A_k$ adódik. Végül a (2) szabály alapján kapjuk a kívánt eredményt: $XY \to B$. □

Az 5.4. lemma bizonyítása. Legyen $\mathcal{F} \vdash_A X \to Z$. Megmutatjuk, hogy

$$\mathcal{F}' \vdash_U \mathcal{G} = \{X \to B \mid B \in Z \setminus X\}.$$

Ehhez készítünk egy U rendszerbeli levezetést, párhuzamosan $X \to Z$ levezetésével az Armstrong-rendszerben. Tegyük fel, hogy az $\langle f_1, \cdots, f_m \rangle$ sorozat $X \to Z$ egy levezetése. Az új levezetést indukcióval készítjük el, biztosítva, hogy minden $f_i = V \to W$-re a hozzá tartozó

$$\mathcal{F}'_i = \{V \to C \mid C \in W \setminus V\}$$

halmaz minden eleme szerepeljen az új levezetésben, még mielőtt f_{i+1}-re kerülne a sor. Ez azt jelenti, hogy minden i-re létezik az új levezetésnek egy $\langle f'_1, \cdots, f'_j \rangle$ prefixe, melyre

$$\bigcup_{p=1}^{i} \mathcal{F}'_p = \bigcup_{q=1}^{j} \{f'_q\}.$$

Adott i-re a levezetés ezen a tulajdonságát $\mathcal{P}_i$-vel jelöljük. Vegyük észre, hogy ez a halmaz $\mathbb{D}_c^+$ része, így ha csak az (U) szabályt alkalmazzuk, a levezetés érvényes lesz az U rendszerben $\mathbb{D}_c^+$ fölött. Az indukció indításához az új levezetés kezdőprefixe az üres sorozat.

Legyen $0 < i \le m$ és tegyük fel, hogy a régi levezetés $\langle f_1, \cdots, f_i \rangle$ prefixe alapján elkészült az új levezetés $\mathcal{P}_i$ feltételnek megfelelő $\langle f'_1, \cdots, f'_j \rangle$ prefixe. Az indukciós lépés az új levezetés prefixének (opcionális) bővítését jelenti. A bővített prefix valamely $h \ge 0$-ra $\langle f'_1, \cdots, f'_j, f'_{j+1}, ..., f'_{j+h} \rangle$ lesz, melyre

$$\mathcal{F}'_{i+1} \subseteq \bigcup_{q=1}^{j+h} \{f'_q\}$$

és így a $\mathcal{P}_{i+1}$ tulajdonság teljesül. Ha a helyes bővítési lépést megmutatjuk, kész a bizonyítás, mert m-re az állítás indukcióval adódik, és

$$f_m = X \to Z, \qquad \mathcal{F}'_m = \mathcal{G}.$$

A bővítési lépés az eredeti bizonyításban szereplő f_{i+1} feldolgozását jelenti. f_{i+1} négyféle lehet: vagy $\mathcal{F}$-beli, vagy a (0) axióma szerint került be, vagy pedig az (1), illetve (2) szabályokkal a levezetésben már korábban szerepelt függőség(ek)ből képzett. Tekintsük át a négy esetet.

Tegyük fel, hogy $f_{i+1} \in \mathcal{F}$. Ez esetben $\mathcal{F}'_{i+1}$ elemeit egyszerűen hozzáírjuk az új bizonyításhoz. Mivel $\mathcal{F}'_i \subseteq \mathcal{F}'$, a lépés helyes.

Ha f_{i+1}-et a (0) axióma „példányosításával” kaptuk, akkor $F'_{i+1} = \emptyset$, és emiatt $\mathcal{P}_{i+1}$ automatikusan teljesül az új levezetés bővítése nélkül.

Tegyük fel, hogy $f_{i+1} = X'V'W' \to Y'V'$-t az (1)-es szabállyal kaptuk valamely előző $f_{i'} = X' \to Y'$ alapján. Ez esetben az

$$\mathcal{F}'_{i+1} = \{X'V'W' \to D \mid D \in Y'V' \setminus X'V'W'\}$$

halmaz elemeivel kell bővíteni az új levezetést (hacsak nem üres). $\mathcal{F}'_{i+1}$ valamely $X'V'W' \to D$ elemére

$$D \in Y'V' \setminus X'V'W' = Y' \setminus X'V'W' \subseteq Y' \setminus X',$$

és így $X' \to D \in \mathcal{F}'_{i'}$, ami már korábban be kellett kerüljön az új levezetésbe. Ha az (U) szabályt $k = 0$-val erre alkalmazzuk, éppen $X'V'W' \to D$-t kapunk, amelyet így hozzáveszünk az új levezetéshez. Ezt $\mathcal{F}'_{i+1}$ minden egyes elemére megtesszük, biztosítva $\mathcal{P}_{i+1}$ teljesülését.

A hátralévő eset az, amikor $f_{i+1} = X' \to Z'$ a (2)-es szabállyal származik valamely korábbi két

$$f_{i_1} = X' \to Y' \quad \text{és} \quad f_{i_2} = Y' \to Z'$$

függőségből. Ha $\mathcal{F}'_{i+1} \neq \emptyset$, akkor minden $X' \to E \in \mathcal{F}'_{i+1}$ függőséget hozzá kell vegyünk az új levezetéshez valamely korábbiak következményeként. Így válik igazzá $\mathcal{P}_{i+1}$. $\mathcal{F}'_{i+1}$ egy ilyen elemére $E \in Z' \setminus X'$. Ha

$$E \in (Z' \setminus C) \cap Y' \subseteq X' \setminus Y',$$

akkor $X' \to E \in \mathcal{F}'_{i_1}$, és mivel ezt már f_{i_1} feldolgozásakor hozzáadtuk az új levezetéshez, nem kell tennünk semmit. Különben pedig $E \in Z' \setminus X'Y' \subseteq Z' \setminus Y'$ és $Y' \to E$-t f_{i_2} feldolgozásakor hozzáadtuk az új levezetéshez. Legyen

$$X'' = X' \setminus Y',\ Y'' = Y' \cap X' \quad \text{és} \quad \{A_s \mid 1 \leq s \leq k\} = Y' \setminus X'$$

úgy, hogy $|Y' \setminus X'| = k$ ($k = 0$ is lehetséges, ekkor nincs egy A_s se). Ezekkel a jelölésekkel

$$Y' \to E = Y''A_1 \cdots A_k \to E \quad \text{és} \quad \forall s \in [1..k] : X' \to A_s = X''Y'' \to A_s \in \mathcal{F}'_{i_1}.$$

Miután $\mathcal{F}'_{i_1}$ elemei benne vannak az új levezetésben, (U) alkalmazható az alábbi módon: X'' felel meg az (U) szabály eredeti jelölésében lévő X-nek, Y'' pedig Y-nak. Így éppen az

$$X''Y'' \to E = X' \to E$$

eredményt kapjuk, melyet hozzáadunk az új levezetéshez. A bővítési lépést mind a négy esetre megadtuk, ezzel a lemmát bizonyítottuk. □

Az 5.5. lemma bizonyítása. $V \to A_1 \cdots A_k$ levezethető a (8) szabály többszöri alkalmazásával, kiindulva a $V \to A_1, \cdots, V \to A_k$ függőségekből. Tekintsük ezt első lépésnek minden esetben (a megfelelő V-vel, amely lehet XY, X vagy XZW).

Elsőként (E1) helyességét mutatjuk meg. A $k = 0$ (és így $l = 0$) eset a (0) axiómával és a (3) szabállyal egyszerűen látható. Legyen most $k > 0$, és tegyük fel, hogy

$$XY \to A_1 \cdots A_k, \text{ and } XY \nrightarrow ZA_1 \cdots A_l$$

igaz (azaz eleme a függőségi rendszernek) valamely $0 \leq l \leq k$ esetén. Az (1) szabállyal $XY \to YA_1 \cdots A_k$ adódik, majd a (3) szabály szerint $YA_1 \cdots A_k \nrightarrow ZA_1 \cdots A_l$ következik. Végül az (5) szabállyal megkapjuk az $YA_1 \cdots A_k \nrightarrow Z$ negált függőséget.

Az (E2) a (6) szabályból azonnal következik (a $k = 0$ érdektelen).

Az (E3) megmutatásához először tegyük fel, hogy $k > 0$ és $XZW \to A_1 \cdots A_k$, valamint $XY \nrightarrow ZA_1 \cdots A_k$ igaz ($ZW \neq \emptyset$). Az első függőséget bővítjük az (1) szabállyal, így kapjuk, hogy $XZW \to ZA_1 \cdots A_k$, melyre alkalmazva a (7) szabályt $XY \nrightarrow YZW$ adódik. A kívánt $XY \nrightarrow ZW$ az (5) szabály alapján következik. A $k = 0$ eset megfelel a (4) szabálynak, amely a (7) speciális eseteként ($Z \subseteq Y$) is felfogható.

Mivel – mint éppen láttuk – az UE rendszer szabályai helyesek, minden véges $\mathcal{F} \subset \mathbb{D}_c^+ \cup \mathbb{E}^+$ esetén $\mathcal{F} \vdash_{UE} \neg(\mathrm{E}\Box) \Longrightarrow \mathcal{F} \vdash_{EA} \neg(\mathrm{E}\Box)$. Ebből már látszik, hogy $\mathcal{F}$ ellentmondásos, de teljesen formálisan is megkaphatjuk: pl. az $X \to A_1$ és $X \nrightarrow A_1$ a (6) szabály többszöri alkalmazásával érhető el $\neg(\mathrm{E}\Box)$-ből. □

Az 5.6. lemma bizonyítása. Legyen $\mathcal{F} \vdash_{EA} X \nrightarrow Z$, $Z \setminus X \neq \emptyset$, és tegyük fel, hogy $\mathcal{F}$ ellentmondásmentes. Vegyük észre, hogy ekkor a hozzá tartozó $\mathcal{F}'$ is ellentmondásmentes, mivel ekvivalens $\mathcal{F}$-fel a kiterjesztett Armstrong-rendszer szerint. Megmutatjuk, hogy a $\mathcal{H} = \{X \to U \mid U \subseteq Z \setminus X, U \neq \emptyset\}$ halmaz legalább egy f'' elemére $\mathcal{F}' \vdash_{UE} f''$ teljesül. Ehhez készítünk egy levezetést az UE rendszerben anélkül, hogy alkalmaznánk (E3)-at $k = 0$-val. Ez igazolja a második állítást. A teljesség ezután egyszerűen következik, ha a levezetést még egy lépéssel bővítjük, nevezetesen (E3)-at $k = 0$-val alkalmazzuk, s így $\mathcal{F}' \vdash_{UE} X \nrightarrow Z$.

Az 5.4. lemmához hasonlóan $\mathcal{H}$ egy elemét levezetjük az UE rendszerben, párhuzamosan $X \nrightarrow Z$ egy meglévő levezetésével, mely az Armstrong-rendszerben van felírva. Tegyük fel, hogy ez a meglévő levezetés az $\langle f_1, \cdots, f_m \rangle$ sorozat, melyben az egyes f_i-k pozitív vagy negatív funkcionális függőségek lehetnek. Az új levezetést az 5.4. lemma bizonyítása során tárgyalt módszert bővítve készítjük el. A pozitív függőségek levezetése teljes egészében az ott tárgyalt módon történik, s emlékeztetünk, hogy így az eredeti levezetés egy f_{i+1} függőségének feldolgozásakor már az

$$\mathcal{F}'_i = \{V \to C \mid C \in W \setminus V\}$$

halmaz elemei bekerültek az új levezetésbe, ha $f_i = V \to W$ alakú volt. Minden negált $f_i = V \nrightarrow W$ függőségre igaz, hogy $W \setminus V \neq \emptyset$ (mivel $\mathcal{F}$ ellentmondásmentes[12]), így az

$$\mathcal{F}_i'' = \{V \to U \mid U \subseteq W \setminus V, U \neq \emptyset\}$$

halmaz nem üres. Az új levezetést úgy készítjük el, hogy $\mathcal{F}_i''$ legalább egy eleme szerepeljen benne mire f_{i+1} sorra kerül. Ezen tulajdonság formalizálásához eddigi definícióinkat kiterjesztjük: $\mathcal{F}_i'' = \emptyset$ lesz a pozitív függőségekre és $\mathcal{F}_i' = \emptyset$ a negáltakra. Az invariáns ezáltal így írható fel: minden i-re van az új levezetésnek egy $\langle f_1', \cdots, f_j' \rangle$ prefixe, melyre teljesül, hogy

$$\bigcup_{p=1}^{i} \mathcal{F}_p' \subseteq \bigcup_{q=1}^{j} \{f_q'\}, \quad \bigcup_{p=1}^{i} (\mathcal{F}_p' \cup \mathcal{F}_p'') \supseteq \bigcup_{q=1}^{j} \{f_q'\}$$

és

$$\forall p \in [1..i] : F_p'' \neq \emptyset \Longrightarrow \mathcal{F}_p'' \cap \bigcup_{q=1}^{j} \{f_q'\} \neq \emptyset.$$

Jelöljük az invariáns tulajdonságot $\mathcal{P}_i$-vel. Mivel

$$\bigcup_{q=1}^{j} \{f_q'\} \subset \mathbb{D}_c^+ \cup \mathbb{E}^+$$

igaz, az új levezetés érvényes lesz az UE rendszerben $\mathbb{D}_c^+ \cup \mathbb{E}^+$ felett, ha csaz az UE szabályait alkalmazzuk. Kezdetben az új levezetés az üres sorozat.

Legyen $0 < i \leq m$, és indukcióval tegyük fel, hogy az $\langle f_1', \cdots, f_j' \rangle$ sorozat az új levezetés már elkészített része (prefixe), mely a meglévő levezetés $\langle f_1, \cdots, f_i \rangle$ prefixe alapján készült és $\mathcal{P}_i$ igaz rá vonatkozóan. Az $f_{i+1} \in \mathbb{D}$ eseteket ($\mathcal{F}$ egy pozitív eleme, illetve a (0) példányosítása vagy az (1), illetve (2) szabállyal kapott függőség) az 5.4. lemma bizonyításakor tárgyaltuk. Ha $f_{i+1} \in \mathbb{E}$, akkor $\mathcal{F}_{i+1}'' \neq \emptyset$, és (opcionálisan) bővítjük az új levezetést az $\langle f_1', \cdots, f_j', f_{j+1}', ..., f_{j+h}' \rangle$ sorozattá úgy, hogy

$$\mathcal{F}_{i+1}'' \cap \bigcup_{q=1}^{j+h} \{f_q'\} \neq \emptyset.$$

Ez igazolja az állítást, hiszen $f_m = X \not\to Z$, $\mathcal{F}_m'' = \mathcal{H}$ és m indukcióval érhető el.

Legyen $f_{i+1} \in \mathcal{F} \cap \mathbb{E}$, $f_i = X' \not\to Z'$. Ez esetben $X' \not\to Z' \setminus X' \in F_i''$, és ezzel a negált függőséggel bővítjük az új levezetést, hogy $\mathcal{P}_{i+1}$ teljesüljön.

A második esetben tegyük fel, hogy

$$f_{i+1} = Y' \not\to Z' \, (Z' \setminus Y' \neq \emptyset)$$

a (3) szabállyal valamely előző

$$f_{i_1} = X' \to Y' \quad \text{és} \quad f_{i_2} = X' \not\to Z' \, (Z' \setminus X' \neq \emptyset)$$

[12]Egy önellentmondásos negált függőség ($XY \not\to X$) a kiterjesztett Armstrong-rendszerrel akkor és csak akkor vezethető le, ha a kezdeti függőségi halmaz ellentmondásos. A „csak akkor” irány triviális. A másik irány a (3) szabály $Y = Z$ esetéből, a (6) szabály $Y = \emptyset$ esetéből, illetve a (7) szabály $X = Y$ esetéből egyaránt következik.

függőségek alapján következik. Ekkor az új levezetést valamely $Y' \not\twoheadrightarrow V'$ függőséggel kell bővítenünk egy nem üres $V' \subseteq Z' \setminus Y'$ halmazzal. Vegyünk egy

$$X' \not\twoheadrightarrow U' \in F''_{i_2} \qquad (U' \subseteq Z' \setminus X')$$

olyan függőséget, amely már előfordul az új levezetés eddig elkészített prefixében. Legyen

$$X'' = X' \setminus Y', \qquad Y'' = X' \cap Y',$$
$$Z'' = U' \setminus Y' \subseteq Z' \setminus Y', \qquad \{A_1, \cdots, A_k\} = Y' \setminus X',$$
$$|Y' \setminus X'| = k$$

($k = 0$ is megengedett, ekkor nincsenek A_s-k), és tegyük fel, hogy $A_1, \cdots, A_l$ éppen $U' \cap Y'$ elemei, mely a $Y' \setminus X'$ részhalmaza. Így

$$Z''A_1 \cdots A_l = U' \quad \text{és} \quad X''Y'' \not\twoheadrightarrow Z''A_1 \cdots A_l = X' \not\twoheadrightarrow U'.$$

Minden $s \in [1..k]$-ra

$$X' \to A_s = X''Y'' \to A_s \in \mathcal{F}'_{i_1},$$

és emiatt már akkor bekerültek az új levezetésbe, amikor f_{i_1}-et dolgoztuk fel. Az (E1) szabály alkalmazásával, X''-et írva X helyére, Y''-t írva Y helyére és Z''-t írva Z helyére,

$$Y''A_1 \cdots A_k \not\twoheadrightarrow Z'' = Y' \not\twoheadrightarrow Z''$$

adódik, ha $Z'' \neq \emptyset$. Az új levezetést bővítjük ezzel a negált függőséggel, és $Z'' \subseteq Z' \setminus Y'$ biztosítja, hogy a kívánt $\mathcal{P}_{i+1}$ tulajdonság igaz legyen. Megmutatjuk, hogy Z'' nem lehet üres. Indirekt módon tegyük fel, hogy az. Ekkor

$$X' \not\twoheadrightarrow U' = X' \not\twoheadrightarrow A_1 \cdots A_l.$$

Mivel korábban minden $X' \to A_s$ le lett vezetve az UE rendszerben, $\neg$(E$\square$)-t kapjuk, ami azt jelenti, hogy F' ellentmondásos (ld. 5.5. lemma, az UE-axiómarendszer helyes). Ez viszont ellentmond a feltételnek, miszerint F' ellentmondásmentes, ezért $Z'' \neq \emptyset$.

Ha

$$f_{i+1} = X' \not\twoheadrightarrow Y'Z' \, (Y'Z' \setminus X' \neq \emptyset)$$

a (4) szabállyal következett egy előző $f_{i'} = X' \not\twoheadrightarrow Y'$ alapján, akkor az indukció miatt egy

$$X' \not\twoheadrightarrow U' \in F''_{i'} \, (U' \subseteq Y' \setminus X', \, U' \neq \emptyset)$$

függőségnek már szerepelnie kell az új levezetésben. Mivel $U' \subseteq Y' \setminus X' \subseteq Y'Z' \setminus X'$, $F''_{i'} \subseteq F''_{i+1}$, és így a $\mathcal{P}_{i+1}$ tulajdonság a levezetés bővítése nélkül igaz lesz.

A következő esetben tegyük fel, hogy az

$$f_{i+1} = X'Z' \not\twoheadrightarrow Y' \, (Y' \setminus X'Z' \neq \emptyset)$$

függőség az eredeti levezetésbe az (5) szabály alapján a korábbi $f_{i'} = X'Y' \nrightarrow Y'Z'$ közvetlen következményeként került be. Az indukció miatt az új levezetésben már szerepelnie kell egy

$$X'Z' \nrightarrow U' \in F''_{i'}\,(U' \subseteq Y'Z' \setminus X'Z',\, U' \neq \emptyset)$$

függőségnek és így újfent nincs szükség bővítésre a $\mathcal{P}_{i+1}$ tulajdonság teljesüléséhez, ugyanis $Y'Z' \setminus X'Z' = Y' \setminus X'Z'$, és így $F''_{i'} = F''_{i+1}$.

Most tegyük fel, hogy

$$f_{i+1} = X' \nrightarrow Y'\,(Y' \setminus X' \neq \emptyset),$$

és ez a (6) szabály $f_{i_1} = X' \to Z'$-re és $f_{i_2} = X' \;\; Y'Z'$-re történt alkalmazásával került az eredeti levezetésbe. $\mathcal{P}_{i+1}$ teljesüléséhez egy

$$X' \nrightarrow V' \in F''_{i+1}\,(V' \subseteq Y' \setminus X',\, V' \neq \emptyset)$$

függőséget kell az új levezetésbe beillesztenünk. Az indukciós feltevés miatt $X' \nrightarrow U' \in F''_{i_2}$ már szerepel az új levezetésben egy nem üres $U' \subseteq Y'Z' \setminus X'$ halmazzal, és nem kell tennünk semmit, ha $U' \subseteq Y' \setminus X'$. Legyen

$$\{A_1, \cdots\; A_k\} = (Z' \setminus X') \cap U'$$

különböző A_s-ekkel ($k = 0$ megengedett), és $Z'' = U' \setminus Z'$. Így

$$X' \nrightarrow U' = X' \nrightarrow Z''A_1 \cdots A_k.$$

Mivel minden $X' \to A_s \in F'_{i_1}$, ezek már le vannak vezetve az új bizonyításban. Az, hogy $Z'' \neq \emptyset$, a (3) szabály eseténél tárgyalt módon bizonyítható itt is. Emiatt az (E2) szabály alkalmazható Z'' helyettesítésével Z-be, és az eredmény,

$$X' \nrightarrow Z'' = X' \nrightarrow U' \setminus Z'\,(U' \setminus Z' \subseteq Y' \setminus X')$$

hozzáírható az új levezetéshez, teljesítve $\mathcal{P}_{i+1}$-t.

Az utolsó eset az, amikor az $f_{i+1} = X' \nrightarrow Y'\,(Y' \setminus X' \neq \emptyset)$ a (7) szabállyal valamely korábban levezetett $f_{i_1} = Y' \to Z'$ és $f_{i_2} = X' \nrightarrow Z'\,(Z' \setminus X' \neq \emptyset)$ alapján következett. Ekkor, hogy $\mathcal{P}_{i+1}$ igaz legyen, biztosítanunk kell, hogy valamely

$$X' \nrightarrow V' \in \mathcal{F}''_{i+1}\,(V' \subseteq Y' \setminus X',\, V' \neq \emptyset)$$

bekerüljön az új levezetésbe. Az indukció miatt egy

$$X' \nrightarrow U' \in F''_{i_2}\,(U' \subseteq Z' \setminus X',\, U' \neq \emptyset)$$

már bekerült az új levezetésbe. Ha $U' \setminus Y' = \emptyset$, akkor $U' \subseteq Y' \setminus X'$, és P_{i_2}-ből következik P_{i+1}. Ellenkező esetben legyen

$$\{A_1, \cdots A_k\} = U' \setminus Y', \qquad |U' \setminus Y'| = k \quad (k > 0),$$

és

$$Z'' = U' \cap Y', \qquad Y'' = X' \cap Y',$$
$$X'' = X' \setminus Y', \qquad W'' = Y' \setminus X'U'.$$

Ezekkel a jelölésekkel $U' = Z''A_1 \cdots A_k$, $X' = X'' \setminus Y''$ és $Y' = Y''Z''W''$. Minden $s \in [1..k]$-ra

$$Y' \to A_s = Y''Z''W'' \to A_s \in F'_{i_1},$$

és ezek mindegyike bekerült az új levezetésbe f_{i_1} feldolgozásakor. Alkalmazzuk az (E3) szabályt $k > 0$-ra, X''-et X helyére, Y''-t Y helyére, Z''-t Z helyére és W''-t W helyére írva az imént említett függőségekre az

$$X' \nrightarrow U' = X''Y'' \nrightarrow Z''A_1 \cdots A_k$$

függőséggel együtt, melynek eredménye $X''Y'' \nrightarrow Z''W''$ lesz. Vegyük észre, hogy $Y'' \cap Z'' = X' \cap Y' \cap U'$ és ez üres, mivel $U' \subseteq Z' \setminus X'$. Emiatt

$$U' \cap Y' \subseteq Y' \setminus X',$$

és

$$Z''W'' = (U' \cap Y') \cup (Y' \setminus X'U') = Y' \setminus X' \neq \emptyset.$$

Így az (E3) szabállyal levezetett függőséget az új levezetéshez hozzáírva $\mathcal{P}_{i+1}$ teljesülni fog.

Ezzel az összes lehetséges esetet kimerítettük, így az új levezetés konstrukciója a régi minden egyes f_{i+1} lépése alapján elvégezhető; és ha az eredeti levezetés a kiterjesztett Armstrong-rendszerben $X \nrightarrow Z$-t bizonyította, akkor az új az UE rendszerben $\mathcal{H}$ egyik elemét eredményezi. Ez a lemma első állítását igazolja. Vegyük észre, hogy az (E3) szabályt nem alkalmaztuk $k = 0$-val egyik esetben sem, így a lemma második állítását is igazoltuk.

Tekintsük a harmadik, utolsó állítást, amikor is $\mathcal{F}$ ellentmondásos, de nem tartalmaz redundáns negált függőségeket. Tegyük fel, hogy valamely X és Y attribútumhalmazra $\mathcal{F} \vdash_{EA} X \to Y$ és $\mathcal{F} \vdash_{EA} X \nrightarrow Y$ is igaz.

Legyen $\{C_1, \cdots, C_k\} = Y \setminus X$. Mivel a (3)-(7) szabályokkal nem vezethető le pozitív függőség, $\mathcal{F} \vdash_A X \to Y$ kell teljesüljön, és így $\mathcal{F}' \vdash_U \{X \to C_i\}$ minden $i \in [1..k]$-ra az 5.4. lemma szerint. Megmutatjuk, hogy vagy $\mathcal{F}' \vdash_{UE} X \nrightarrow C_1 \cdots C_l$ igaz valamely $l \in [1..k]$-ra, vagy $\neg(\mathrm{E}\Box)$ (esetleg már attribútumokkal) levezethető. Kezdjük el $X \nrightarrow C_1 \cdots C_k$ levezetését az UE rendszerben, párhuzamosan $X \nrightarrow Y$ bizonyításával csakúgy, mint az ellentmondásmentes esetben. Ha megnézzük a módszert, mindaddig működik, amíg az eredeti levezetésben egy $X'Y' \nrightarrow X'$ alakú önellentmondásos függőséghez érkezünk. Ha most egy pillanatra feltesszük, hogy az UE rendszerben a negált függőségek jobb oldala lehet üres, az új levezetést tovább építhetnénk, mely azt eredményezné, hogy $X'Y' \nrightarrow \emptyset$-zal kell bővítenünk az (E1), (E2) vagy (E3) szabály alkalmazásának következtében. Ahelyett, hogy elvégeznénk ezt a bővítést, könnyen észrevehetjük, hogy ennek a lépésnek az előfeltétele (az (E1) és (E2) esetén $Z = \emptyset$, (E3)-ra pedig $ZW = \emptyset$) éppen $\neg(\mathrm{E}\Box)$ valamely $A_1, \cdots, A_{k'}$ attribútumokkal. Így arra a következtetésre jutunk, hogy még mielőtt elakadna az új levezetés elkészítése, $\neg(\mathrm{E}\Box)$ már le kellett vezetve legyen. Ha nem akad el, akkor pedig $X \nrightarrow C_1 \cdots C_k$ levezethető valamely $l \in [1..k]$-ra, ami megint csak $\neg(\mathrm{E}\Box)$-nak felel meg. Vegyük észre, hogy ismét nem alkalmaztuk (E3)-at $k = 0$-ra ahhoz, hogy megkapjuk $\neg(\mathrm{E}\Box)$-t. $\Box$

Az 5.7. lemma bizonyítása. Az (S) és (T) szabályok az (U) speciális esetei: $k = 0$, $|X| = 1$ illetve $k = 1$, $|X| = 0$.

Legyen $k > 0$ és $|X| > 0$. Megmutatjuk, hogy minden $0 \leq t \leq k$-ra,

$$\{XY \to A_1, \cdots, XY \to A_k, YA_1 \cdots A_k \to B\} \vdash_{ST} XYA_1 \cdots A_t \to B$$

igaz. Az igazolandó állítás éppen a $t = 0$ eset, $t = k$-ra pedig az (S) alkalmazásával egyszerűen következik. Tegyük fel t-re vonatkozó indukcióval, hogy

$$XYA_1 \cdots A_{t+1} \to B$$

levezethető. Ekkor $XY \to A_{t+1}$-ből az (S) szabállyal $XYA_1 \cdots A_t \to A_{t+1}$ adódik, és ebből (T)-vel azonnal következik

$$XYA_1 \cdots A_t \to B.$$

A $t = 0$ esetet indukcióval kapjuk. □

Az 5.8. lemma bizonyítása. Az $\{(U)\}$ és az $\{(S), (T)\}$ ekvivalenciáját az 5.7. lemma állította és az imént bizonyítottuk. Maradnak tehát a negált függőségekre vonatkozó szabályok. Az (NS) és az (NT1) az (E1) speciális esetei, $|X| = 1$, $k = 0\,(l = 0)$, ill. $|X| = 0$, $k = 1$, $l = 0$ választással. Hasonlóan, (NT2) és $(*)$ az (E3) speciális esetei rendre $k = 1$ és $k = 0$ választással.

(NT2S) egyszerűen szimulálható az $Y \to B$ függőség (S)-sel $YZ \to B$-re történő bővítésével, majd az (NT2) alkalmazásával $W = \emptyset$ esetén $Y \not\to Z$-t kapunk.

$\neg$(E□)-ből indulva $\neg$(□) levezethető (NT2S) ismételt alkalmazásával. A fordított eset triviális, mivel (□) az (E□) speciális esete ($k = 1$).

Tekintsük az (E1) $k = 0$, $l = 0$, $|X| > 1$ esetét. Ezt könnyen szimulálhatjuk (NS) ismételt alkalmazásával. Legyen $k > 0$. Ha $l > 0$, akkor (NT2S)-sel az A_i-ket ($1 \leq i \leq l$) a $XY \not\to ZA_1 \cdots A_l$ jobb oldaláról lépésenként elvehetjük (az $XY \to A_i$ függőség igaz minden i-re). (E1) még nem tárgyalt esete a $k > 0$, $l = 0$. (NT1) alkalmazásával $XYA_1 \not\to Z$ kapható. Minden $XY \to A_j$ ($1 < j \leq k$) függőséget $XYA_1 \cdots A_{j-1} \to A_j$-re bővítünk (S)-sel, hogy (NT1) ismételten alkalmazható legyen. Így minden j-re megkapjuk $XYA_1 \cdots A_j \not\to Z$-t. Amint $j = k$ lesz, (NS)-sel elvesszük X-et a bal oldalról és megkapjuk a kívánt $YA_1 \cdots A_k \not\to Z$ függőséget.

Az (E2) szabály csak $k > 0$-ra releváns és $X \not\to ZA_1 \cdots A_j$ levezethető (NT2S)-sel lépésenként minden $j \in [0..k]$-ra. $j = 0$-ra éppen a kívánt $X \not\to Z$ függőséget kapjuk.

Tekintsük (E3)-at $k > 0$-ra. Ehhez

$$XZW \to A_k\text{-t}$$

bővítjük (S)-sel

$$XZA_1 \cdots A_{k-1}W \to A_k\text{-ra.}$$

Ezután (N2)-vel $XY \nrightarrow ZA_1 \cdots A_{k-1}W$ adódik ($ZA_1 \cdots A_{k-1}$ helyettesítendő Z helyére az eredeti képletben). Az $A_1 \cdots A_{k-1}$ attribútumok eliminálhatók a jobb oldalról (NT2S) ismételt alkalmazásával, így $XY \nrightarrow ZW$ adódik.

Az utolsó eset (E3) $k = 0$-val, ami $(*)$ (esetleges ismételt) alkalmazásának felel meg. Ezzel beláttuk az ekvivalenciát. Vegyük észre, hogy a $(*)$ szabályt nem alkalmaztuk más esetben. Ez igazolja a lemma utolsó állítását is, azzal együtt, hogy az NST minden szabálya (beleértve a $(\Box)$ axiómát is) az UE valamely szabályának (vagy (E$\Box$)-nak) speciális esete és $(*)$ az egyetlen, amely megfelel (E3)-nak $k = 0$-ra. $\Box$

Az 5.1. lemma bizonyítása. Triviális: A (P) és a (Q) rendre az (NS) és az (NT1) speciális esetei ($|Z| = 1$, $B = Z$); (R) pedig az (NT2) speciális esete ($Z = \emptyset$, $|W| = 1$, $A = W$). $\Box$

Az 5.2. lemma bizonyítása. Először tegyük fel, hogy $\mathcal{F}$ ellentmondásmentes (az 5.6. és 5.8. lemmák értelmében ez azt jelenti, hogy $\mathcal{F} \nvdash_{NST'} \neg(\Box)$), és hogy $\mathcal{F} \vdash_{NST'} X \nrightarrow A$ igaz. Tekintsük $X \nrightarrow A$ egy levezetését. A szabályok alakjából látható, hogy a levezetés lépései átrendezhetők oly módon, hogy minden pozitív függőség megelőzze a negatív függőségeket. Ha kihagyjuk a levezetés azon elemeit, amelyek nem szükségesek $X \nrightarrow A$ bizonyításához, láthatjuk, hogy csak az első negált függőség kell hogy $\mathcal{F}$-beli legyen, hiszen nincs olyan szabály, melynek az előfeltétele két vagy több negált függőséget tartalmazna. Továbbá, az így megmaradt negált függőségek sorrendje olyan, hogy mindegyik az előzőből lett egy lépésben levezetve valamely szabály alkalmazásával (és esetleg egy pozitív függőség figyelembe vételével). Jelölje $\langle \delta_1, \cdots, \delta_s, \varepsilon_t, \cdots, \varepsilon_1 \rangle$ ezt az NST' rendszerben felírt, szűkített és átrendezett levezetést ($\delta_i \in \mathbb{D}_c^+$, $\varepsilon_j \in \mathbb{E}_c^+$ minden $i \in [1..s]$-re és $j \in [1..t]$-re). Vegyük észre, hogy $\varepsilon_1 = X \nrightarrow A$, és mivel $\varepsilon_t \in \mathcal{F}$, továbbá ez az egyetlen ε_i ezzel a tulajdonsággal, $(\mathcal{F} \cap \mathbb{D}_c^+) \cup \{\varepsilon_t\} \vdash_{NST'} X \nrightarrow A$ igaz. Készítünk egy ekvivalens levezetést a PQRST rendszerben az eredetivel párhuzamosan. A pozitív függőségeket ($\langle \delta_1, \cdots, \delta_s \rangle$) változatlanul átírjuk az új bizonyítás elejére, mivel az (S) és a (T) szabályok közösek az NST' és a PQRST szabályrendszerben. A negatív függőségeket ($\langle \varepsilon_t, \cdots, \varepsilon_1 \rangle$) egyenként tekintve lecseréljük egy $\langle \delta_{s+1}, \cdots \delta_r, \zeta_u, \cdots, \zeta_1 \rangle$ sorozatra ($r \geq s$, $u \in \mathbb{N}^+$; $\delta_i \in \mathbb{D}_c^+$, $\zeta_j \in \mathbb{E}_c^+$ minden $i \in [s+1..r]$-re és $j \in [1..u]$-ra), mely érvényes levezetés lesz a PQRST rendszerben $\mathbb{D}_c^+ \cup \mathbb{E}_c^+$ fölött. A konstrukció során biztosítjuk, hogy minden $i \in [1..t]$-re az új levezetésnek legyen egy $\langle \delta_1, \cdots, \delta_{h_i} \rangle$ prefixe és egy $\langle \zeta_{j_i}, \cdots, \zeta_1 \rangle$ szuffixe ($h_i \geq h_{i-1} \geq s$ és $j_i \geq j_{i-1}$ minden $i \in [2..t]$-re) úgy, hogy $\langle \delta_1, \cdots, \delta_{h_i}, \zeta_{j_i}, \cdots, \zeta_1 \rangle$ egy érvényes levezetés, mely igazolja az

$$(\mathcal{F} \cap \mathbb{D}_c^+) \cup \{\varepsilon_i\} \vdash_{PQRST} X \nrightarrow A$$

állítást. Továbbá biztosítjuk, hogy $\zeta_{j_i} = V \nrightarrow A$ minden $\varepsilon_i = V \nrightarrow W$-re, és $VA \rightarrow E \in \{\delta_1, \cdots, \delta_{h_i}\}$ minden $E \in W \setminus A$-ra. E két tulajdonságot jelölje $\mathcal{Q}_i$ adott $i \in [1..t]$-re. Első lépésként $\mathcal{Q}_1$ a $\zeta_1 = \varepsilon_1$ választással, $h_1 = s$ és $j_1 = 1$ mellett igaz. t pedig indukcióval érhető el, az alább részletezett módon eljárva minden lépésben. Ha $\zeta_{j_t} = \varepsilon_t$, akkor $u = j_t$, és a bizonyítás kész. Máskülönben

$$\zeta_{j_t} = Y \nrightarrow A, \qquad \varepsilon_t = Y \nrightarrow E\,(A \neq E),$$

és ekkor $\mathcal{Q}_t$ szerint $YA \to E$ szerepel az új levezetés prefixében. Ekkor $\zeta_{j_t+1} = \varepsilon_t$-t kell az elkészített szuffix elejére illeszteni (legyen $u = j_t + 1$), mivel ζ_{j_t} levezethető az (R) szabállyal az $YA \to E$ és az ε_t függőségekből kiindulva. Ez az (R) egyetlen szükséges alkalmazása, mint azt az alábbiakban látni fogjuk.

Indukcióval tegyük fel, hogy a PQRST-beli új levezetés egy $\langle \delta_1, \cdots, \delta_{h_i} \rangle$ prefixe és egy $\langle \zeta_{j_i}, \cdots, \zeta_1 \rangle$ szuffixe már elkészült az eredeti levezetés pozitív függőségeinek és $\langle \varepsilon_i, \cdots, \varepsilon_1 \rangle$ szuffixének feldolgozásával ($i \in [1..t-1]$) úgy, hogy teljesíti $\mathcal{Q}_i$-t. Az indukciós lépésben végzünk egy (opcionális) bővítést a prefixen, a szuffixon (vagy mindkettőn), hogy $\mathcal{Q}_{i+1}$-et is teljesítse. Többféle eset lehetséges aszerint, hogy mely szabály alkalmazásával kaptuk ε_i-t ε_{i+1}-ből az eredeti levezetésben ($\varepsilon_i \notin \mathcal{F}$ mint láttuk).

Tegyük fel, hogy $\varepsilon_i = Y \nrightarrow Z$ és $\varepsilon_{i+1} = YC \nrightarrow Z$ (az (NS) szabályt alkalmazta az eredeti levezetés), és $\zeta_{j_i} = Y \nrightarrow A$. Legyen

$$j_{i+1} = j_i + 1, \qquad \zeta_{j_{i+1}} = YC \nrightarrow A$$

(ez érvényes lépés, hiszen ζ_{j_i} levezethető (P) alkalmazásával $\zeta_{j_{i+1}}$-re). Hogy $\mathcal{Q}_{i+1}$ teljesüljön, bővítjük az új levezetés prefixét minden olyan

$$YCA \to E\,(E \in Z \setminus A)$$

függőséggel, amely még nem szerepel benne. Ezek az indukció alapján már a bizonyításban szereplő $YA \to E$-ből (S) alkalmazásával állíthatók elő. Vegyük észre, hogy $A \neq C$ mivel $\mathcal{F}$ ellentmondásmentes ($A = C$ azt jelentené, hogy $YC \to E$ igaz minden $E \in Z \setminus C = Z$-re, és ez ellentmond $YC \nrightarrow Z$-vel).

A második eset az, amikor az (NT1) szabályt alkalmazta az eredeti levezetés $\varepsilon_{i+1} = Y \nrightarrow Z$-re, és valamely $\delta_k = Y \to A'\,(k \leq s)$, hogy megkapja

$$\varepsilon_i = YA' \nrightarrow Z\text{-t.}$$

Az indukciós feltevés miatt

$$\zeta_{j_i} = YA' \nrightarrow A \qquad \text{és} \qquad \forall E \in Z \setminus A\, \exists l \in [1..h_i] : \delta_l = YA'A \to E.$$

A szuffixet bővíthetjük $\zeta_{j_{i+1}} = \zeta_{j_i+1} = Y \nrightarrow A$-val, mivel ζ_{j_i} a (Q) szabállyal egy lépésben megkapható δ_k-ból és ebből a $\zeta_{j_{i+1}}$-ből. Továbbá $YA \to E$ levezethető minden $E \in Z \setminus A$-ra (S)-sel és (T)-vel, kiindulva $Y \to A'$-ből és $YA'A \to E$-ből, melyek már szerepelnek az új levezetésben. A prefixet bővítjük az $YA \to A'$ és az $YA \to E$ függőségekkel, ha azok még nem szerepelnek benne, így megkapjuk $\mathcal{Q}_{i+1}$-et.

A harmadik és egyben utolsó eset az (NT2) szabályé. Legyen $\varepsilon_i = Y \nrightarrow ZW$, $\varepsilon_{i+1} = Y \nrightarrow ZB$, és k olyan, hogy $\delta_k = YZW \to B$ ($k \leq s$). Az indukció miatt

$$\zeta_{j_i} = Y \nrightarrow A, \qquad \text{és} \qquad \forall E \in ZW \setminus A\, \exists l \in [1..h_i] : \delta_l = YA \to E.$$

Legyen $j_{i+1} = j_i$ (a szuffixot nem bővítjük). Q_{i+1} teljesüléséhez még $YA \to E$-t kell hozzávenni a prefixhez, ha az nem szerepel még benne, minden $E \in ZB \setminus A$-ra. Mivel minden $E \in Z \setminus A$-ra már szerepel benne, csak az $YA \to B$-vel kell bővíteni, ha egyáltalán kell (ez esetben, $A \neq B$). Az (U) szabállyal az $YA \to B$ megkapható a fent említett δ_k és δ_l függőségekből. Az 5.7. lemma szerint az (U) kifejezhető az (S) és a (T) szabályokkal és ennek a bizonyításnak a lépéseivel kell bővítenünk az új levezetést.

Egy ellentmondásmentes függőségi halmazra a PQRST-beli új levezetés elkészíthető a fenti lépésekkel. Ha viszont a kezdeti $\mathcal{F}$ halmaz ellentmondásos, akkor $\neg(\Box)$ levezethető az NST' rendszerben. Egy ilyen levezetés átalakítása elakadhat, ha tartalmaz egy olyan $\varepsilon_k = Y \nrightarrow Z$ függőséget, amelyre $k > 1$ és $\mathcal{F} \vdash_{ST} Y \to E$ minden $E \in Z$-re (ld. az (NS) szabály esetét). Az első ilyen tulajdonságú ε_k-t kiválasztva (a lehető legnagyobb k indexszel) és az eredeti levezetést itt elvágva (csak az ez előtti függőségig tekintve) befejezhető az átalakítás. Először tetszőlegesen kiválasztunk egy $D \in Z$ attribútumot $\zeta_1 = Y \nrightarrow D$-hez és $Y \to E$ levezetésének lépéseit minden $E \in Z$-re hozzáadjuk a $\langle \delta_1, \cdots, \delta_s \rangle$ prefixhez, majd tovább folytatjuk az átalakítás lépéseit a tárgyalt módon. Az eljárás nem akad el, és mind $Y \to D$, mind pedig $Y \nrightarrow D$ szerepelni fog a levezetésben, jelezve, hogy $\mathcal{F}$ ellentmondásos. $\Box$

Az 5.1. tétel (i) állításának bizonyítása. Az ST rendszer helyessége és teljessége az Armstrong-rendszer helyességéből és teljességéből, valamint az 5.3., 5.4. és 5.7. lemmákból következik. Vegyük észre, hogy az 5.4. lemmában, ha $\mathcal{F} \in \mathbb{D}_c^+$ és $X \to Z \in \mathbb{D}_c^+$, akkor $\mathcal{F} = \mathcal{F}'$, és pontosan egy $B \in Z \setminus X$ attribútum és egy $X \to Z = X \to B$ függőség van. Így egyszerűen azt kapjuk, hogy

$$\mathcal{F} \vdash_{ST} \delta \iff \mathcal{F} \vdash_U \delta \iff \mathcal{F} \vdash_A \delta \iff \mathcal{F} \models \delta.$$

Az 5.1. tétel (ii) állításának bizonyítása. Megmutatjuk, hogy (S) (T) utáni alkalmazása szimulálható (S) kétszeri, majd (T) kétszeri alkalmazásával. Tegyük fel, hogy a kiindulási függőségek $YA \to B$ és $Y \to A$. (T), majd (S) alkalmazásával $Y \to B$ és $YC \to B$ adódik. E helyett (S)-et alkalmazva mindkét kezdeti függőségre $YCA \to B$-t és $YC \to A$-t kapunk. Erre a két függőségre alkalmazva (T)-t, megkapjuk $YC \to B$-t, és $Y \to B$ levezethető (T)-vel ezt követően a két kiindulási függőségből. $\Box$

Az 5.3. tétel (i) állításának bizonyítása. Az NST rendszer helyessége és teljessége (a $(\Box)$ nélkül) a kiterjesztett Armstrong-axiómarendszer helyességéből és teljességéből, valamint az 5.5., az 5.6. és az 5.8. lemmából következik. Vegyük észre, hogy az 5.6. lemmában, ha

$$\mathcal{F} \subset \mathbb{D}_c^+ \cup \mathbb{E}^+ \qquad \text{és} \qquad X \nrightarrow Z \in \mathbb{E}^+,$$

akkor

$$\mathcal{F} = \mathcal{F}' \qquad \text{és} \qquad X \nrightarrow Z = X \nrightarrow Z \setminus X \quad (Z \setminus X \neq \emptyset).$$

Az 5.1. tétel (i) állításának bizonyításához hasonlóan kapjuk, hogy

$$\mathcal{F} \vdash_{NST} \delta \iff \mathcal{F} \vdash_{UE} \delta \iff \mathcal{F} \vdash_{EA} \delta \iff \mathcal{F} \models \delta$$

minden ellentmondásmentes $\mathcal{F}$-re. A helyesség érvényes függetlenül attól, hogy $\mathcal{F}$ ellentmondásmentes-e, így $\mathcal{F} \vdash_{UE} (\Box) \Longrightarrow \mathcal{F} \vdash_{EA} (\Box)$. A fordított irány ellentmondásos esetben az 5.6. lemmából következik, mivel $F = F' \subset \mathbb{D}_c^+ \cup \mathbb{E}^+$.

Az 5.3. tétel (ii) állításának bizonyítása. A helyesség $\mathbb{D}_c^+ \cup \mathbb{E}^+$ fölött és a teljesség $\mathbb{D}_c^+$ fölött e tétel (i) állításának triviális következménye. Az (i) állításhoz hasonlóan $\mathcal{F}' = \mathcal{F}$. Az 5.6. és az 5.8. lemma alapján (azok utolsó állításaira hivatkozva) az alábbiakat kapjuk, ha $\mathcal{F}$ ellentmondásmentes:

$$\begin{aligned} \mathcal{F} \Vdash X \not\twoheadrightarrow Y \in \mathbb{E}^+ &\Longrightarrow F \vdash_{UE} X \not\twoheadrightarrow Y \Longrightarrow \\ &\Longrightarrow [\exists Z \subseteq Y,\, Z \neq \emptyset : \mathcal{F} \vdash_{UE} X \not\twoheadrightarrow Z \text{ (E3) nélkül } k = 0\text{-ra}] \Longrightarrow \\ &\Longrightarrow [\mathcal{F} \vdash_{NST} X \not\twoheadrightarrow Z \text{ ugyanarra a } Z\text{-re } (*) \text{ nélkül}]. \end{aligned}$$

Ellentmondásos $\mathcal{F}$-re pedig azt kapjuk, hogy

$$[\mathcal{F} \vdash_{UE} \neg(\mathrm{E}\Box) \text{ (E3) nélkül } k = 0\text{-ra}] \Longrightarrow [\mathcal{F} \vdash_{NST} \neg(\Box)(*) \text{ nélkül}]. \qquad \Box$$

Az 5.2. tétel (i) állításának bizonyítása. Az állítás az 5.3. tételből és az 5.1., 5.2. lemmákból triviálisan következik.

Az 5.2. tétel (ii) állításának bizonyítása. Az (i) állítás szerint minden pozitív következmény levezethető úgy, hogy először (S)-et alkalmazzuk amíg lehet, majd utána (T)-t amíg lehet. Világos, hogy a (P), (Q) és (R) szabályok alkalmazása maradhat ezek utánra. Azt kell belátnunk, hogy az (R) szabály (P) vagy (Q) utáni alkalmazása helyettesíthető egy olyan levezetéssel, amelyben (R) megelőzi (P) és (Q) alkalmazásának összes előfordulását.

Először tegyük fel, hogy (R)-et közvetlenül (Q) után alkalmaztuk. Az $Y \rightarrow A$, $Y \not\twoheadrightarrow B$ és $YAC \rightarrow B$ függőségekből indulva (Q)-val $YA \not\twoheadrightarrow B$-t kapunk, majd ezt követően (R) alapján $YA \not\twoheadrightarrow C$-t. $YA \not\twoheadrightarrow C$ egy alternatív levezetéséhez vegyük észre, hogy $YC \rightarrow B$-t már le kellett vezetnünk, mivel az egy pozitív függőség, mely (T)-vel következik. Az (R) szabályt alkalmazva először $YC \rightarrow B$-re és $Y \not\twoheadrightarrow B$-re, $Y \not\twoheadrightarrow C$-t kapunk, és a kívánt $YA \not\twoheadrightarrow C$ függőség a (Q) szabály alapján következik. Ezután a (Q) szabályt alkalmazzuk ismét az eredeti levezetésnek megfelelően, így kapjuk a $YA \not\twoheadrightarrow B$ függőséget.

A másik eset az, amikor az (R) szabályt a (P) után alkalmazzuk: az $YC \not\twoheadrightarrow B$ és az $YA \rightarrow B$ függőségekből kiindulva az $Y \not\twoheadrightarrow B$ a (P) szabály alapján következik, majd $Y \not\twoheadrightarrow A$-t kaphatjuk az (R)-rel. Vegyük észre, hogy $YAC \rightarrow B$-t előzőleg már le kellett vezetnünk az (S) szabállyal. Ebből és az $YC \not\twoheadrightarrow B$-ből (R)-rel $YC \not\twoheadrightarrow A$ következik, és az $Y \not\twoheadrightarrow A$ (P)-vel megkapható. (P) második alkalmazásával az eredetivel egyező módon ezután megkaphatjuk a $Y \not\twoheadrightarrow B$-t.

Mindezekből következik, hogy nem veszítünk következtetési lehetőséget azzal, ha (R) mindig megelőzi (P) és (Q) alkalmazását. Az pedig triviális, hogy (P) és (Q) sorrendje tetszőleges.

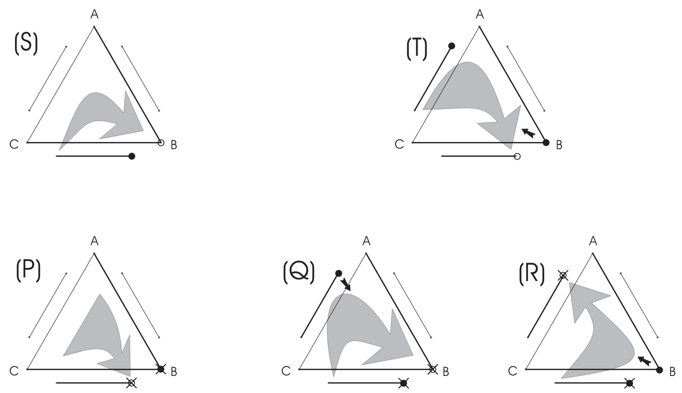
Ha $|\mathcal{G}| = 1$, akkor a levezetés átalakítható az 5.2. lemma bizonyításában szereplő módon, mivel a PQRST szabályai speciális esetei az NST' szabályainak (ld. az 5.1. lemma bizonyítását), és így minden PQRST-beli levezetés NST'-beliként is felfogható. Az átalakított levezetésben (R) alkalmazására csak egyszer kerül sor. $\square$

6. Következtetés grafikus úton

Implikált függőségek levezetése a függőségi rendszer diagramján közvetlenül is elvégezhető valamely axiómarendszer szabályainak grafikus megfelelőivel. Erre a legalkalmasabbnak a PQRST rendszer tűnik, a szabályok egyszerűsége miatt (csak egy halmazparamétert tartalmaznak, Y-t, és azon kívül két attribútumparamétert, B-t és A-t vagy C-t). Általánosságban a PQRST rendszerben egy levezetési lépés egy magasabb dimenziós alakzat (pl. háromszög) és egy nálánál eggyel kisebb dimenziós „határoló" alakzat (pl. a háromszög mellé rajzolt szakasz) egy-egy csúcspontjának kölcsönhatásaként tekinthető.

6.1. Grafikus szabályok

Elsőként az $n = 3$-as esetet részletezzük.



5. ábra. Az (S), (T) és (P), (Q), (R) szabályok grafikus (háromszöges) megfelelői.

A szabályoknak megfelelő grafikus minták az 5. ábrán láthatók (háromattribútumos eset, $Y = \{C\}$). A kis fekete nyilak az implikációhoz szükséges kontextust (a *segédfüggőség*eket) jelölik, míg a nagy szürke nyilak az implikáció irányát, hatását. Az (S) szabály egy egyszerű bővítést jelöl (a bal oldal bővül egy attribútummal, így eggyel nagyobb dimenziós implikált függőség jön létre), míg a (T) szabály „forgatásnak" (rotáció) vagy „szűkítésnek" (redukció) tekinthető. A forgatás azt jelenti, hogy a függőség jobb oldala kicserélhető (egy megfelelő, eggyel magasabb dimenziós függőség megléte esetén), míg a szűkítés egy attribútum elhagyását jelenti a bal oldalról (a bal oldal attribútumain teljesülnie kell hozzá a megfelelő függőségnek). A két interpretáció között a lényegi különbség az, hogy mit tekintünk az implikáció kiinduló függőségének és mit a szükséges kontextusnak. Mi a forgatásos megközelítést használjuk.

A negált függőségek bővítési szabálya a (Q), forgatási szabálya pedig az (R). Mindkettő pozitív segédfüggőséggel működik. E két szabály a (T) negáltjának is tekinthető. A (P) szűkítési (redukciós) szabály, a (Q)-val ellentétes hatású, de segédfüggőség nélküli. A (P) az (S) negáltjának felel meg.

A 3. ábrán jól követhető a grafikus szabályok alkalmazása. A bal szélső diagramon az (S) szabályt alkalmaztuk, míg a jobb szélsőn a (P)-t kétszer. A középső egy tranzitivitási esetet mutat be, melyet az (S) szabály kétszeri, majd ezt követően a (T) egyszeri alkalmazásával zárunk le.

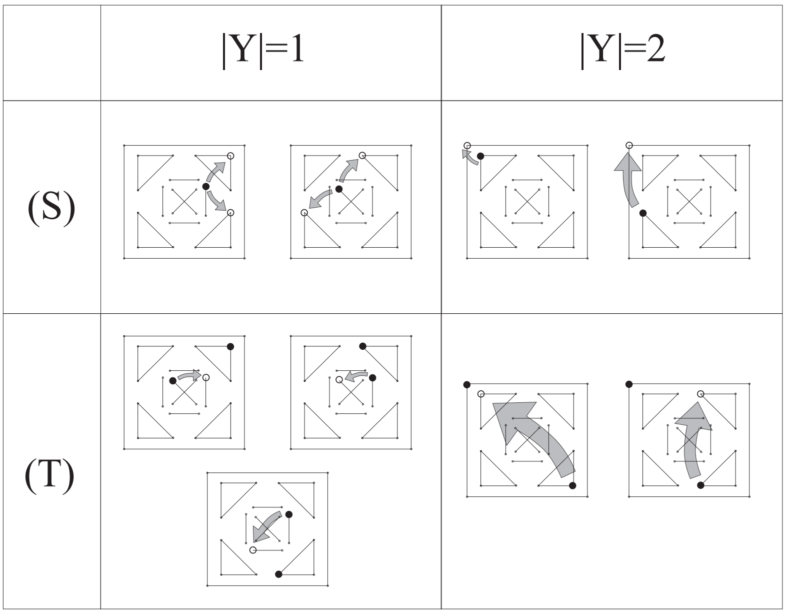
E grafikus szabályminták megadhatók több attribútum esetére is. Négy attribútum esetén a szabályokban szereplő Y szerint már kétféle eset lehet, az ugyanis tartalmazhat egy vagy két attribútumot. A tetraéderes reprezentációban pl. az (S) már felrajzolt háromszöges mintája mellett megjelenik egy eggyel magasabb dimenziós minta: háromszögcsúcs bővítése tetraédercsúccsá. A (T) szabály második mintája pedig egy háromszög forgatása megfelelő tetraédercsúcs megléte mellett. A 6. ábrán az (S) és a (T) négyzetes reprezentációra vonatkozó mintái láthatók. A tetraéder négyzetre történt cseréjével a szimmetria megbomlott, ezért egy esetnek két vagy három minta felel meg.

A négyattribútumos esetre a 7. ábrán láthatunk példát: hogyan szimulálható a tranzitivitás az (S) és (T) szabályokkal a $\{C \to BD, BD \to A\} \vdash C \to A$ esetben. A kezdő diagramhoz elsőként a $C \to BD$ függőséget szét kell bontani $\{C \to B, C \to D\}$-re. Ezután a következtetés például az alábbi sorrendben végezhető el: 1. $BD \to A \vdash_{(S)} BCD \to A$; 2. $C \to B \vdash_{(S)} CD \to B$; 3. $CD \to B$ (kontextus:) $BCD \to A \vdash_{(T)} CD \to A$; 4. $C \to D$ (kontextus:) $CD \to A \vdash_{(T)} C \to A$.

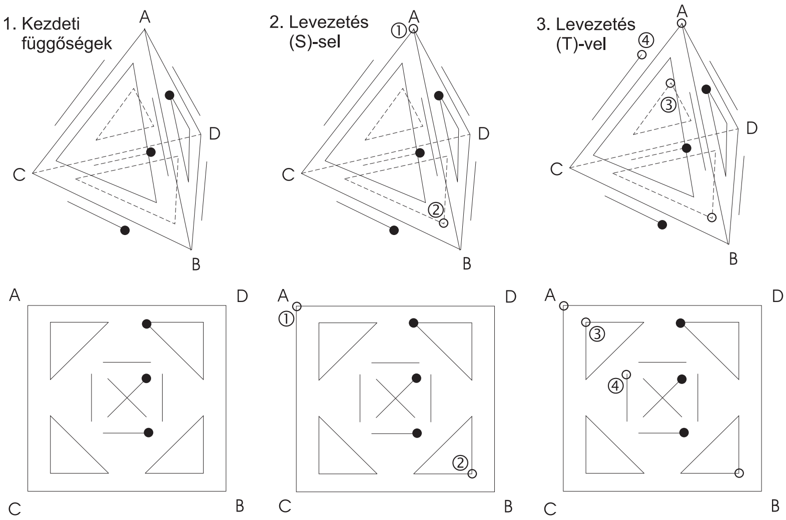
A fenti esetben a tranzitivitásig eljutottunk, de a halmaz még nem zárt. A 4. ábra egy zárt halmazt mutat be.

6.2. Az ST és az STRPQ algoritmus

Az egyik leggyakrabban felmerülő kérdés a függőségi halmazok kezelésekor (és a velük való sématervezéskor) az, hogy a kezdetben felvett függőségek (tele körök) milyen más függőségeket implikálnak, azaz milyen típusú kapcsolatot definiáltunk a kezdeti függőségekkel. Ha negált függőségeink is vannak, az is elképzelhető, hogy ellentmondásos esetet adtunk meg. Ezeket a kérdéseket a függőségi halmaz lezárásával válaszolhatjuk meg, melyet a (PQR)ST axiomatizációval szisztematikusan is elvégezhetünk.



6. ábra. Az (S) és a (T) szabály mintái a négyzetes reprezentációban.



7. ábra. Következtetési példa a tetraéderes, illetve a négyzetes reprezentációban: tranzitivitás (a számok egy lehetséges implikációs sorrendet jelölnek).

Az 5. fejezet 5.1. tétele (ii) állítása szerint teljes levezetési módszert jelent a pozitív függőségekre, ha először az (S) szabályt alkalmazzuk, amíg lehet, majd utána a (T)-t. Ezt a módszert *ST algoritmus*nak nevezzük.[13] Ha negált függőségeink is vannak, a (P), a (Q) és az (R) szabályokkal mindezt bővítve (az 5.2. tétel (ii) része alapján) az alábbi, ún. *STRPQ* módszert kapjuk az összes implikált (pozitív és negatív) függőség szisztematikus levezetéséhez:

1. Az inputként megadott nemredundáns, szingleton függőségekből álló halmazzal kezdve,
2. bővítjük a függőségek bal oldalait az (S) szabály alapján, az összes lehetséges módon, majd
3. alkalmazzuk a (T) szabályt amíg lehet, ezután
4. alkalmazzuk az (R) szabályt amíg lehet, majd
5. szűkítjük és bővítjük a negált függőségek bal oldalait a (P) és a (Q) szabály alapján, az összes lehetséges módon.
6. Az output az így keletkezett függőségi halmaz lesz.

A fenti algoritmus akár közvetlenül az input függőségi rendszer diagramján is elvégezhető. Az általánosított háromszöges reprezentáció alapján elvileg adatszerkezet is tervezhető a függőségi rendszerekhez, melyen az algoritmus természetes módon futtatható.

Az ST módszer még konkrétabbá tehető a dimenziók figyelembe vételével. Először kezdjünk az (S) lehető legalacsonyabb dimenziós mintáival, majd folyamatosan haladjunk egyre magasabb dimenziók felé. A (T)-vel pedig éppen fordítva haladjunk. Ha már beláttuk az ST módszer teljességét, könnyen látható, hogy ez esetben is teljes módszert kapunk: egy magasabb dimenziós (S)-minta alkalmazása nem generál nálánál alacsonyabb dimenziós esetet, melyben (S) alkalmazható lenne. Hasonlóan a (T) alkalmazása nem generál olyan új esetet, mely (T) egy magasabb dimenziós mintáját tenné alkalmazhatóvá. Ha ilyen eset létezik, az már a szabály alkalmazása előtt is fennállt.

7. Konklúzió

Annak jegyében, hogy a funkcionális függőségi rendszereket egységként és ne csupán a függőségek összességeként kezeljük, bevezettük a diagramos reprezentációkat (három attribútmra háromszöges, négyre tetraéderes vagy négyzetes, ötre hasonlóan általánosítható). Segítségükkel egyszerűbbé, áttekinthetővé válik a függőségi rendszerek kezelése, megkönnyítve a grafikus adatbázis-tervezést.

[13] Ez a módszer adta az alapját a PROLOG programnak, amellyel a lehetséges függőségi rendszereket generáltuk és összeszámoltuk (ld. 3. fejezet).

A függőségek egyszerűsített formalizmusa (miszerint a redundánsokkal, köztük a triviálisokkal, valamint a nem kanonikusokkal (nemszingletonokkal) nem foglalkozunk) és az ehhez megalkotott ST és PQRST szabályrendszer megszünteti a függőségek hagyományos rendszeréből adódó redundanciákat. A szabályok a grafikus reprezentációhoz is jól illeszkednek: megadhatók nekik megfelelő, diagramon alkalmazható minták, így kis sémákra a grafikus következtetés könnyen, akár kézzel is elvégezhető. Sőt, megadható olyan szabályalkalmazási sorrend, mely teljes, szisztematikus módszert ad egy halmaz lezárására. Ez leírható az (S)*(T)*(R)*((P)*|(Q)*) reguláris kifejezéssel. Az axiomatizáció helyességét és teljességét a nemredundáns, szingleton függőségekre nézve az 5.1. és 5.2. tétel (i) állításai mondják ki, az említett sorrendiséget pedig ezen tételek (ii) állításai.

Nagyobb sémák grafikus reprezentációjára más módszert kell keresni, folyamatban van az itt bemutatottaknak a zárt attribútumhalmazok hálójával [20, 21, 12] való összevetése. Szükséges lehet további dekompozíciós, szeparációs lehetőségek ezen keretek közti vizsgálata is. Szintén érdekes probléma más típusú függőségek hasonló reprezentációjának megadása.

Az (S) és (T) szabályok alapján kis sémákra (legföljebb 5 attribútum esetén) meghatároztuk a lehetséges zárt függőségi halmazok számát. Nagyobb sémákra ($n \geq 6$) ennek általánosítása egyelőre megoldatlan, csupán becslések vannak [10, 21].

Köszönet

Ezúton köszönjük Ásványi Tibor segítségét a PROLOG program hatékonyságának növelésében és Regéci Zoltán Csaba közreműködését a program futtatásában. Ezen túlmenően hálásak vagyunk Molnár Andrea és Benczúr András hasznos megjegyzéseiért, észrevételeiért.

Hivatkozások

[1] S. Abiteboul and R. Hull and V. Vianu: *Foundations of databases.* Addison-Wesley, 1995.

[2] W. W. Armstrong: *Dependency structures of data base relationships.* In J. L. Rosenfeld, editor, *Information Processing 74, Proceedings of IFIP Congress 74*, Stockholm, Aug. 5-10, 1974. North-Holland, Amszterdam. (pages 580–583)

[3] P. Atzeni and V. De Antonellis: *Relational database theory.* Addison-Wesley, 1993.

[4] Catriel Beeri and Martin Dowd and Ronald Fagin and Richard Statman: *On the Structure of Armstrong Relations for Functional Dependencies.* J. ACM 31(1):30–46, 1984.

[5] J. Biskup: *Boyce-Codd normal forma and object normal forms.* Information Processing Letters, 32(1):29–33, 1989.

[6] J. Biskup: *Foundations of information systems.* Vieweg, Wiesbaden, 1995. (német nyelvű)

[7] J. Biskup and J. Demetrovics and L. O. Libkin and M. Muchnik: *On relational database schemes having a unique minimal key.* J. of Information Processing, 27:217–225, 1991.

[8] J. Biskup and T. Polle: *Decomposition of database classes under path functional dependencies and onto contraints.* In Proc. FoIKS'2000, LNCS 1762, pages 31–49. Springer, 2000.

[9] P. De Bra and J. Paredaens: *Horizontal decompositions and their impact on query solving.* SIGMOD Rec., 13(1):46–50, 1982.

[10] G. Burosch and J. Demetrovics and G. O. H. Katona and D. J. Kleitman and A. A. Sapozhenko: *On the number of databases and closure operations.* TCS 78(2):377–381, 1991.

[11] R. Camps: *From ternary relationship to relational tables: A case against common beliefs.* ACM SIGMOD Record, 31(2), pages 46–49, 2002.

[12] N. Caspard and B. Monjardet: *The lattices of closure systems, closure operators, and implicational systems on a finite set: a survey.* Discrete Applied Mathematics **127**:241–269, 2003.

[13] Chen & Associates, Baton Rouge, LA, USA. *ER-designer reference manual*, 1986–1989.

[14] P. P. Chen: *The entity-relationship model: Toward a unified view of data.* ACM TODS 1(1):9–36, 1976.

[15] P. P. Chen, editor: *Proc. 1st Int. ER Conf., ER'79: Entity-Relationship Approach to Systems Analysis and Design.* Los Angeles, USA, 1979. 1980, North-Holland, Amsterdam.

[16] E. F. Codd: *A relational model for large shared data banks.* CACM, 13(6):197–204, 1970.

[17] J. Demetrovics and N. X. Huy: *Structure of Closure in Relational Databases.* In Conference on intelligent management systems, pages 148–154. Bulgarian Academy of Sciences, Várna, 1989.

[18] J. Demetrovics and N. X. Huy: *Translations of relation schemes and representations of closed sets.* PU.M.A.Ser. A, 1(3–4):299–315, 1990.

[19] J. Demetrovics and G. O. H. Katona: *Combinatorial problems of database models.* In Colloquia Mathematica Societatis Janos Bolyai 42, Algebra, Combinatorics and Logic in Computer Science, pages 331–352, Győr, 1983.

[20] J. Demetrovics and L. Libkin and I. B. Muchnik: *Functional Dependencies in Relational Databases: A Lattice Point of View.* Discrete Applied Mathematics, 40(2):155–185, 1992.

[21] J. Demetrovics and L. O. Libkin and I. B. Muchnik: *Functional dependencies and the semilattice of closed classes.* In Proc. MFDBS'89, LNCS 364, pages 136–147, 1989.

[22] J. Demetrovics and A. Molnar and B. Thalheim: *Graphical and Spreadsheet Reasoning for Sets of Functional Dependencies.* Technical Report 402, Graphical and Spreadsheet Reasoning for Sets of Functional Dependencies, http://www.informatik.uni-kiel.de/reports/2004/0402.html, 2004.

[23] J. Demetrovics and A. Molnar and B. Thalheim: *Graphical Reasoning for Sets of Functional Dependencies.* In Proceedings of ER 2004, Lecture Notes in Computer Science 3288, pages 166-179. Springer, 2004.

[24] J. Demetrovics and V. D. Thi: *Some problems concerning Armstrong relations of dual schemes and relation schemes in the relational datamodel.* Acta Cybernetica, 11(1–2):35–48, 1993.

[25] M. Klettke: *Akquisition von Integritätsbedingungen in Datenbanken.* DISBIS 51. infix-Verlag, St. Augin, 1998.

[26] Heikki Mannila and Kari-Jouko Räihä: *Small Armstrong relations for database design.* In PODS '85: Proceedings of the fourth ACM SIGACT-SIGMOD symposium on Principles of database systems, pages 245–250, New York, NY, USA, 1985. ACM Press.

[27] C.W. Morris: *Foundations of the Theory of Signs.* In International Encyclopedia of Unified Science, University of Chicago Press, 1955.

[28] V.C. Storey and H.L. Yang and R.C. Goldstein: *Semantic integrity constraints in knowledge-based database design systems.* Data & Knowledge Engineering, 20:1–37, 1996.

[29] B. Thalheim: *Open problems in relational database theory.* Bull. EATCS, 32:336–337, 1987.

[30] B. Thalheim: *Entity-relationship modeling – Foundations of database technology.* Springer, Berlin, 2000. (ld. http://www.informatik.tu-cottbus.de/~thalheim/HERM.htm)

[31] V. D. Thi: *Minimal keys and antikeys.* Acta Cybernetica, 7:361–371, 1986.

[32] J. D. Ullman and J. Widom: *Adatbázisrendszerek – alapvetés.* Panem–Prentice Hall, 1998 (magyar nyelvű változat).

[33] C.-C. Yang: *Relational Databases.* Prentice-Hall, 1986.

(Beérkezett: 2006. május 27.)

DEMETROVICS JÁNOS
MTA SZTAKI
1111 Budapest, Kende u. 13-17.
Email: demetrovics@sztaki.hu

MOLNÁR ANDRÁS
ELTE IK, Információs Rendszerek Tanszék
1117 Budapest, Pázmány Péter sétány 1/C
Email: modras@elte.hu

BERNHARD THALHEIM
Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität Kiel
Olshausenstraße 40, 24098 Kiel, Németország
Email: thalheim@is.informatik.uni-kiel.de

GRAPHICAL AXIOMATIZATION OF SETS OF FUNCTIONAL DEPENDENCIES IN RELATIONAL DATABASES

János Demetrovics, András Molnár and Bernhard Thalheim

An important task during relational database schema design is to specify invariant properties, which the database instances must obey. These semantical conditions can be formulated as integrity constraints. The most fundamental type of constraints is functional dependency. Direct usage of its traditional formalism is rare in practice, due to its complexity and redundancies. However, in some cases, a complete and unambiguous specification can only be reached by using functional dependencies. We propose a novel approach for representing sets of functional constraints for small relational schemata. This graphical representation allows management of constraint sets much simpler and surveyable than the traditional notation. It supports reasoning on constraint sets in a natural way by an appropriate, powerful axiomatization and allows easy derivation. The axiomatization has an order of rule application property, yielding an algorithmic method for deriving the full knowledge an initial set of constraints holds. Considering complexity, we have computed the number of different possible dependency sets and types for schemata with at most five attributes.