

Adatbáziskezelés

Egyed-kapcsolat modell

Csima Judit

BME, VIK,
Számítástudományi és Információelméleti Tanszék

2018. szeptember 5.

- Célja: a modellezendő valóságdarabhoz adatbázisséma létrehozása.
- Elvárás: jól írja le a valóságot, könnyű legyen a gyakori kérdéseket és módosításokat megtenni
- Részei:
 - 1 Terv készítése (nagyon fontos rész, ha rossz tervet csinálunk, később nehéz módosítani) valamilyen modellező eszköz/nyelv segítségével (pl. E/K diagram).
 - 2 A terv átalakítása formálisabb leírássá (tipikusan E/K-ból relációs séma megadása).
 - 3 Az adatbázisséma formális megadása a rendszer által kívánt DDL-en (ez az átalakítás már viszonylag automatikusan megy, a DDL persze rendszerfüggő).

Adatmodellezés, mit tanulunk

- Először az első lépéssel foglalkozunk, a tervezéssel: E-K modell (egyed-kapcsolat modell). Ez egy grafikus modell a séma megtervezésére.
- Közben végig lesz majd arról szó, hogy hogyan kell a grafikus tervet átírni relációs sémára: ez még mindig egy rendszertől független, elméleti séma lesz (de már nem grafikus)
- Aztán pedig majd lesz az, hogy az SQL DDL-jével hogyan kell megvalósítani a relációs sémát: ez egy konkrét adatbáziskezelő rendszerben levő megvalósítás, az adattáblák létrehozása és a megkötések rögzítése

Egyed-kapcsolat modell

Adatmodellező eszköz, azaz többé-kevésbé formális jelölésrendszer, adatok és a köztük levő kapcsolatok megadására.

Vannak más modellező eszközök is, de ez a legelterjedtebb.

Angolul E/R, azaz entity-relationship modell

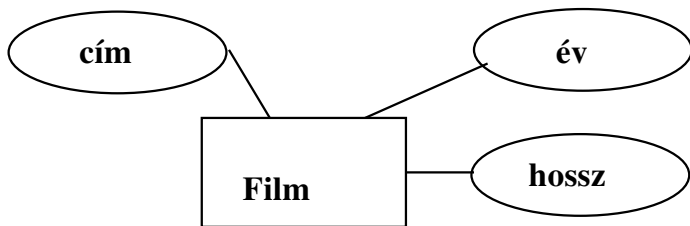
Szemléletes, könnyű vele dolgozni. Egy rajzot készítünk, ez ábrázolja az adatelemeket és a köztük levő kapcsolatot is.

E/K modell, alapfogalmak

- *egyedhalmazok*: ezeket akarjuk tárolni, illetve ezekről dolgokat, (pl. pilóta, utas, járat), elemei (példányai) az egyedek (pl. konkrét pilóták, utasok és járatok)
- az egyedhalmazokhoz *attribútumok* tartoznak, ezeket a tulajdonságokat fogjuk az egyedhalmaz minden egyedéről tárolni (pl. pilóta neve, rangfokozata, fizetése, stb.)
- *kapcsolatok*: az egyedhalmazok közötti kapcsolatok, pl. járat utasai, járat személyzete

Egyedhalmaz attribútumokkal, ábrázolás

Film(cím, gyártási_év, hossz, ...), rajzon:

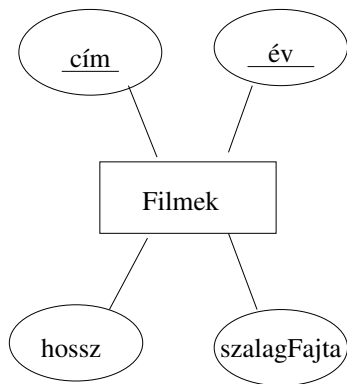


Egyedhalmaz kulcsa

Olyan attribútum vagy attribútumhalmaz, ami az egyedet egyértelműen azonosítja. (Pl. hallgatónál Neptun-kód vagy filmnél (gyártási év, cím) pár).

Egy kulcsot aláhúzással jelölünk (a kulcsba tartozó attribútumokat aláhúzzuk), ha több kulcs is van, akkor azt az ábrán nem lehet jelölni, ezeket szövegesen mellékeljük

E/K modell, kulcs ábrázolása



Film(cím, év, hossz, szalagFajta)

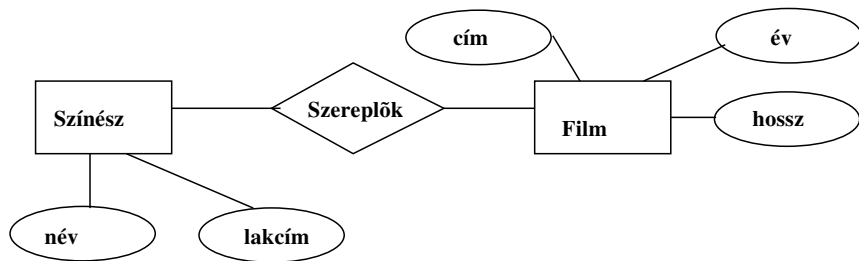
Egyedhalmaz kulcsa

A tervezéskor döntjük el, hogy mik alkossanak kulcsot (persze a valóságot szem előtt tartva). A kulcshoz tartozó attribútumoknak értékeket adva, legfeljebb egy egyed létezhet, amihez ezek az értékek tartoznak.

Néha tûnhet úgy az aktuális adatokból, hogy valami kulcs, mert akkor éppen nincs két egyed ugyanolyan értékekkel, (pl. a film címe kulcsnak tûnhet, amíg nem csinálnak remake-et semmiből), de ettől meg nem lesz kulcs valami, az csak a deklarációtól függ.

Kapcsolatok ábrázolása

Szereplők(Film, Színész):



Kis kitérő: relációs adatmodell

Legfontosabb és leggyakoribb a létező adatmodellek közül.

Most: hogyan kell erre átírni az E/K modellt.

Később:

- 1 alpműveletek, elvi keret
- 2 konkrét nyelv: SQL (sémadefinícióra, adatmódosításra és lekérdezésre)
- 3 tervezés: minél jobb séma kialakítása, matematikai elmélet

Relációs adatmodell

Egyetlen alapfogalom (nincs külön egyedhalmaz és kapcsolat): reláció.

Leginkább úgy gondolunk a relációra, mint egy síkbeli táblázatra:

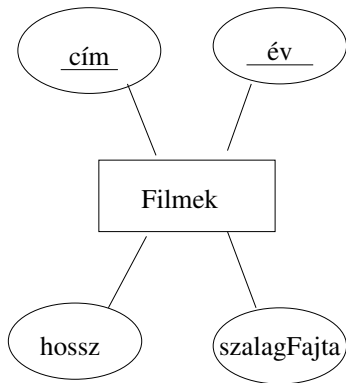
R_1	A_1	A_2
	1	y
	1	z
	3	y

R_2	A_1	A_3
	2	y
	1	z

Itt R_1 a reláció neve, A_1 és A_2 az attribútumok nevei, a sorok pedig a reláció elemei. Az oszlopokban levő értékek az attribútumokhoz tartozó értékészlethez kerülnek ki.

E/K modell átírása relációsra, alapeset

Egyedhalmaz attribútumokkal:

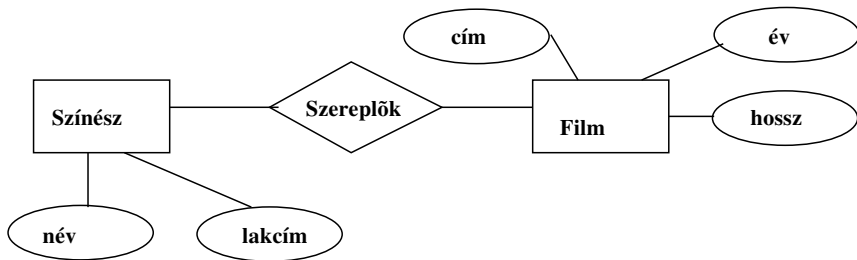


Film(cím, év, hossz, szalagFajta)

A reláció kulcsa = az egyedhalmaz kulcsa

Kapcsolatok ábrázolása

Szereplők(Film, Színész):



Egy kapcsolat példányai

Ha az $R(E_1, E_2, \dots, E_{10})$ kapcsolat 10 egyedhalmazt köt össze, akkor az R kapcsolat egy példánya egy 10 hosszú vektor $(e_1, e_2, \dots, e_{10})$, ahol az e_i egy egyed az E_i egyedhalmazból.

Például Szereplők(Film, Színész) kapcsolat esetén példányok:

(Éhezők viadala, Jennifer Lawrence)

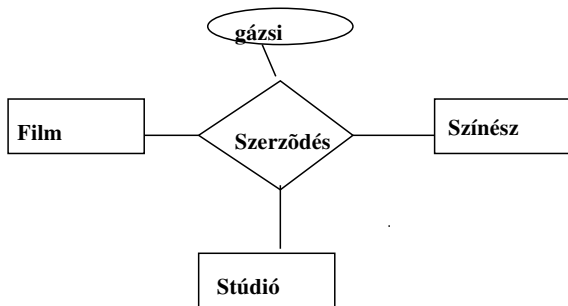
(Éhezők viadala, Elizabeth Banks)

(Napos oldal, Jennifer Lawrence)

(Testről és lélekről, Borbély Alexandra)

Kapcsolat attribútuma

Az E/K modellben a kapcsolatnak is lehet attribútuma (ez nem minden modellben van így):



Itt a gázi a szerződéshez tartozik, ami a filmet, a színészt és a stúdiót köti össze.

Kapcsolatok típusa bináris kapcsolat esetén

- $R(E_1, E_2)$ több-több (many-many) kapcsolat: egy E_1 -beli egyedhez több E_2 -beli tartozhat és egy E_2 -beli egyedhez több E_1 -beli tartozhat, például Szerepel(Film, Színész)
- $R(E_1, E_2)$ E_2 irányba több-egy (many-one) kapcsolat: egy E_1 -beli egyedhez csak egy E_2 -beli tartozhat, de egy E_2 -beli egyedhez több E_1 -beli is tartozhat, például Anyja(Személy, Személy)
- $R(E_1, E_2)$ egy-egy (one-one) kapcsolat: egy E_1 -beli egyedhez csak egy E_2 -beli tartozhat és egy E_2 -beli egyedhez is csak egy E_1 -beli is tartozhat, például Házastársa(Személy, Személy)

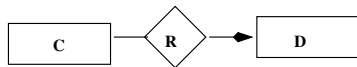
Kapcsolatok típusa

- Fontos: a kapcsolat típusa modellezési kérdés, azt mutatja, hogy mit gondolunk a világról
- Ha egy kapcsolatot valamelyik irányba “egy”-nek jelölünk, akkor az azt jelenti, hogy a relációs modellben majd lesz egy megkötés, ami ezt kikényszeríti.
- Az, hogy egy kapcsolat valamelyik irányba “több”, az csak egy lehetőség: tartozhat egy Színészhez több Film is, de persze lehetnek elsőfilmes színészek is.

Kapcsolatok típusának ábrázolása bináris kapcsolat esetén

Az “egy” irányt nyíl jelzi, azaz ott van nyíl, amelyik egyedhalmazból csak egy tartozhat a másik egyedhalmaz egy egyedéhez.

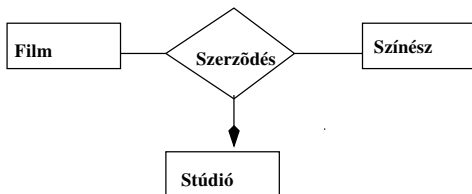
Például, ha a C egyedhalmaz egy egyedéhez csak egy D -beli tartozhat az R kapcsolatnál, de egy D -belihez tartozhat több C -beli is, akkor:



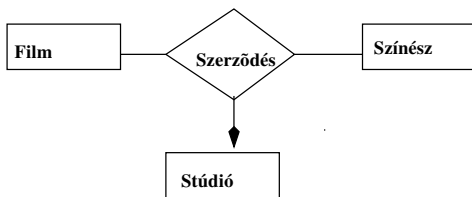
Ha egy-egy kapcsolat van, akkor persze mindkét oldalra kell a nyíl.

Többágú kapcsolat típusa

Az $R(E_1, E_2, \dots, E_n)$ kapcsolat E_i felé egyirányú, ha igaz az, hogy a maradék $E_1, E_2, \dots, E_{i-1}, E_{i+1}, \dots, E_n$ egyedhalmazokból bárhogyan választva ki egy-egy egyed, maximum egy olyan E_i -beli e_i egyed van, amivel az $(e_1, e_2, \dots, e_{i-1}, e_i, e_{i+1}, \dots, e_n)$ egyedvektorra az R kapcsolat fennáll. (Természetesen lehet egy többágú kapcsolat egynél több irányban is "egy" jellegű.) Példa:



Többágú kapcsolatra példa

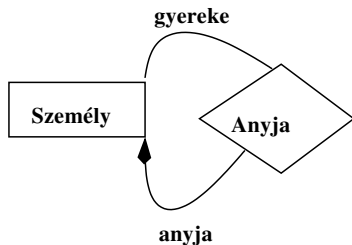


Egy rögzített (film, színész) párhoz csak egy stúdió tartozik, de pl. egy rögzített (stúdió, film) párhoz tartozhat több színész.

Nem minden írható le ilyen módon, de nem baj, úgyis az a fontos, hogy majd a relációs megadásnál pontosak tudjunk lenni.

Kapcsolatok ábrázolása, szerepek

Ha egy kapcsolatban egy egyedhalmaz többször is szerepel, akkor a nyilakon/vonalakon jelöljük a különböző szerepeket. pl:



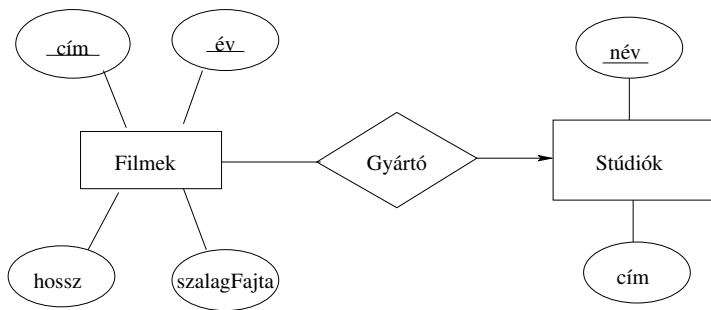
Fontos megcímkézni a vonalakat, mert ahhoz a szerephez tartozik az “egy” nyíl, ami az anyára vonatkozik, a gyerekághoz nem kell.

Kapcsolat átírása relációvá

Minden kapcsolatból külön reláció, attribútumai: a kapcsolatban résztvevő egyedhalmazok kulcsainak uniója + kapcsolat attribútumai (esetleg átnevezés)

Az így kapott reláció kulcsa: a kapcsolatban résztvevő egyedhalmazok kulcsainak uniója

Kapcsolat átírása relációvá



Gyártó(cím, év, stúdióNév)
Film(cím, év, hossz, szalagFajta)
Stúdiók(stúdióNév, cím)

Kapcsolat átírása relációvá, speciális eset

Ha bináris több-egy kapcsolatról van szó, akkor van jobb megoldás is:

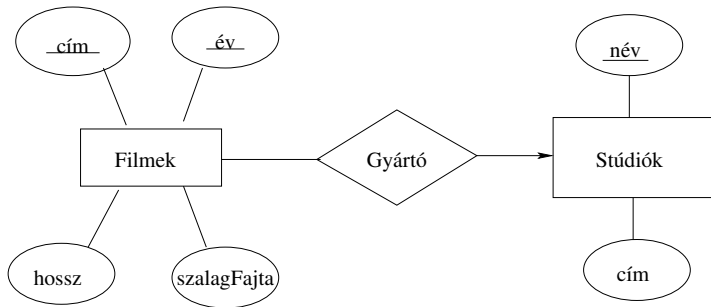
Nem veszünk fel külön relációt a kapcsolatnak, hanem ha az E és F közti kapcsolat F felé egyirányú, akkor az E egyedhalmaz átírásakor bevesszük az F osztály kulcsát is.

Miért jó?

- az E -beli kulcs meghatározza az E -beli egyedet, az pedig meghatározza az F -belit
- eggyel kevesebb tábla lesz
- mivel egy E -belihez csak egy F -beli tartozik, ezért nem lesz redundáns

Az E hez tartozó reláció kulcsa E kulcsával egyezik meg.

Kapcsolat átírása relációvá, speciális eset



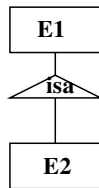
Így most nem kell külön tábla a kapcsolatnak, hanem a Film(cím, év, hossz, szalagFajta, **stúdióNév**) lesz a Film tábla.

Alosztályok

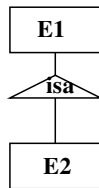
Egy egyedhalmaz speciális tulajdonságú, de egymáshoz hasonló egyedei alkotják. Ezeket érdemes együtt kezelni, de úgy, hogy a (nagyobb) egyedhalmazba való tartozásuk is megmaradjon.

Ehhez egy speciális (*isa*, magyarul *azegy*) kapcsolat van. (Motiváció: A student is a person = a diákok alosztályát alkotják az embereknek)

Jelölés, ha E_2 alosztálya E_1 -nek:



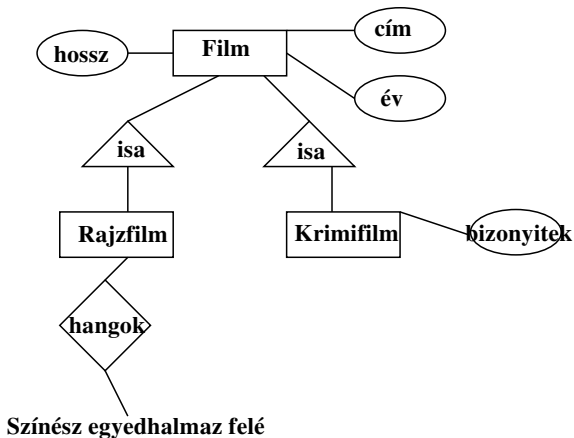
Alosztályok



Az első esetben a nyíl arra mutat, amelyik a fő osztály, ez most nem a több-egy kapcsolatnál megszokott nyíl, az *isa* ilyen szempontból is speciális kapcsolat. A második esetben az alosztály van alul. Az alárendelt halmaz örökli a fő halmaz attribútumait és kapcsolatait, de persze lehetnek neki sajátjai is.

Alosztályokra még egy példa

A Film egyedhalmazon belül akarhatunk külön Rajzfilm, Vígjáték, Krimifilm kategóriákat:



Alosztályokra még egy példa

Ha egy egyed több alosztályba is tartozik, akkor az attribútumait/kapcsolatait a felmenőtől szedi össze. Így az E/K modellben egy olyan film, ami krimi is és rajzfilm is, három helyről szedi össze az attribútumait/kapcsolatait: a címét, hosszát és gyártási évét a Film egyedhalmazból, a hangjait a Rajzfilm alosztályból, a bizonyítékot pedig a Krimifilmből.

A relációs modellre való átíráskor majd ezt ügyesen kezeljük.

Alosztályok kezelése a relációsra való átíráskor

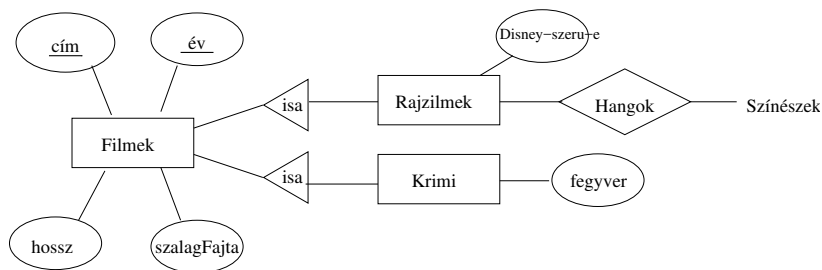
Mivel E/K-ban egy egyed lehet több egyedhalmazban is lehet egyszerre, ezért az egy filmre vonatkozó információk szét vannak szórva. A relációs sémára való átíráskor gondoskodunk róla, hogy a részinfókból vissza tudjuk állítani az egészet (ha kell).

Első megoldás:

Minden alosztályhoz a főosztály kulcsát és saját attribútumait rendeljük. Az alosztály kulcsa a főosztály kulcsa lesz, így a kapcsolataiba is ezt viszi magával az alosztály.

Az "isa" kapcsolathoz nem rendelünk relációt.

Alosztályok kezelése a relációsra való átíráskor



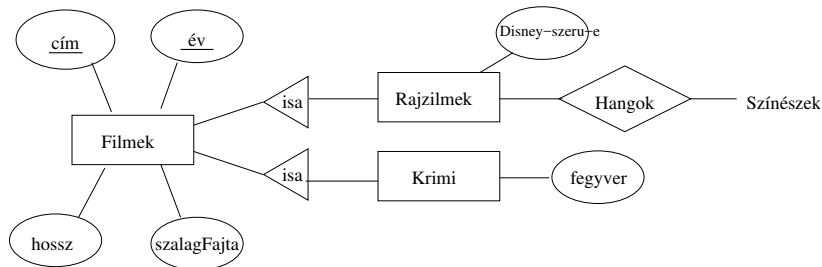
```
Film(cím, év, hossz, szalagFajta),  
Rajzfilm(cím, év, Disney-szeru-e),  
Krimi(cím, év, fegyver),  
Hangok(cím, év, Színésznév)
```


Alosztályok kezelése a relációsra való átíráskor

Ennek a módszernek a hátránya: egy film információi több helyre vannak szórva.

Pl. Macskafogónál: a hossz és a szalagfajta a Film-ben, az, hogy nem Disney-is, az a Rajzfilmben, hangok a Hangokban. De ezeket az infókat össze lehet rakni, a (cím, év) kulcs menti természetes illesztéssel.

Másik megoldás NULL értékkel



Film(cím, év, hossz, szalagFajta, Disney-szerű-e, fegyver)

A hiányzó helyeket NULL-al töltjük ki (pl. NULL van a Disney-szerű-nél ott, ami nem rajzfilm)

Másik megoldás NULL értékkel

Hátrányok:

- 1 elveszíthetünk információt. Pl. egy olyan krimiről, amiben nincs fegyver, nem tudjuk, hogy krimi, illetve ha egy rajzfilmről nem tudjuk, hogy Disney-s vagy sem, akkor elvesztjük azt az infót is, hogy rajzfilm
- 2 NULL értékekre nagyon kell figyelni

NULL értékek, ízelítő

Ez a két SQL kérdés vajon ugyanaz?

```
SELECT * FROM Dolgozó WHERE (fizetés < 400 000 ) OR ( Fizetés ≥  
400 000);
```

```
SELECT * FROM Dolgozó;
```

Nem! Erről majd bővebben SQL-nél, most csak annyi a lényeg, hogy résen kell lenni, ha NULL-ról van szó.

Megszorítások

Olyan megszorításokat is ábrázolni akarunk amik nem fejezhetők ki pusztán az attribútumok és a kapcsolatok felsorolásával vagy a kapcsolatok típusával (pl. több-egy).

Ezek további infók, amik

- a séma részei, ezért már a tervezéskor kell rájuk gondolni,
- olyan megkötéseket tartalmaznak, amikre majd mindig figyelni kell.

Nem világos, hogy mik a jó megkötések, miket lehet jól kezelni.

Tipikus megszorítások

Kulcs: olyan attribútum, vagy attribútumhalmaz megadása, ami az egyedet egyértelműen azonosítja. (Pl. hallgatónál Neptun-kód vagy filmnél (gyártási év, cím, stúdió) hármás).

A tervezéskor döntjük el, hogy mik alkossanak kulcsot (persze a valóságot szem előtt tartva). A kulcshoz tartozó attribútumoknak értékeket adva, legfeljebb egy egyed létezhet, amihez ezek az értékek tartoznak.

Néha tûnhet úgy az aktuális adatokból, hogy valami kulcs (mert akkor éppen nincs két egyed ugyanolyan értékekkel, (pl. a film címe kulcsnak tûnhet, amíg nem csinálnak remake-et semmiből), de ettõl meg nem lesz kulcs valami, az csak a deklarációtól függ.

Tipikus megszorítások

- *Egyértékûségi megszorítások*: elõírhatjuk, hogy valami érték vagy értékkombináció legyen egyedi. Pl. ilyen a kulcsok megadása is, vagy az, hogy bizonyos attribútumok értékei meghatározzák más attribútumok értékeit (ezek lesznek majd a funkcionális függések a relációs modellben, errõl késõbb lesz szó).

Például ilyen az, hogy az irányítószám meghatározza a várost

- *Hivatkozási épség*: a hivatkozott dolognak léteznie kell, pl. ha egy tárgynál szerepel egy hallgató, akkor az a hallgató a hallgatók alapadatait tartalmazó táblában is benne legyen

Tipikus megszorítások

- *Értelmezési tartomány korlátozása*: attribútum lehetséges értékeire megkötés (pl. magasság legyen kisebb 300-nál, film gyártási éve 1800 utáni). Módszerek erre: konkrétan majd az SQL DDL-jénél.
- *Egyéb megszorítások*: minden más, pl. kapcsolat fokának korlátozása (egy filmnek max. 10 szereplőjét akarjuk nyilvántartani).

Megszorítások, általános kérdések

- Alapkérdés: Mik a jó megszorítások? Miket lehet megvalósítani? Ez persze majd a konkrét megvalósítástól függ, a konkrét DDL-től, de azért jó lenne már a modellezéskor is annyit leírni, amennyit csak lehet, aztán majd meglátjuk.
- A megszorítások haszna
 - jobban/valósághoz közelebbi módon le lehet velük írni a világot
 - segíthetik a tárolást (elég pl. a kulcsattribútumokat megadni kereséskor)

Megszorítások ábrázolása E/K modellben

- *Kulcsok*: egy kulcsot aláhúzással jelölünk (a kulcsba tartozó attribútumokat aláhúzzuk), a többi kulcsot az ábrán nem lehet jelölni, ezeket szövegesen mellékeljük
- *Egyértékűség*: kapcsolatnál: nyilakkal jelezhető, ha valamerre “egy” a kapcsolat, egyéb megkötések szövegesen

Megszorítások ábrázolása E/K modellben

- Elsődleges kulcs aláhúzással, többi szövegesen
- Hivatkozási épség: lehet a rajzon jelezni, ha egy kapcsolatnál azt szeretnénk, hogy pontosan egy egyed tartozzon egy kiválasztott egyedhez. Ilyenkor kerek nyilat használunk:



Ebben az esetben minden filmhez pontosan egy stúdiónak kell tartoznia.

Megszorítások ábrázolása E/K modellben

Kapcsolat fokát lehet korlátozni, pl:



Ekkor egy filmhez 10-nél kevesebb színészt rendelünk.

Megszorítások átírása relációsra

Erről majd később lesz szó, amikor a relációs modell elméleti kereteit tanuljuk (funkcionális függések), illetve amikor az SQL DDL-jét tanuljuk.

Most elégedjünk meg annyival, hogy a kulcsokat át tudjuk valahogyan írni, a többről meg azt képzeljük, hogy minden, az E/K digramon jelölt megszorítás szövegesen van leírva és majd meglátjuk, hogy hogyan lehet ezeket implementálni.