

Introducció a l'enginyeria del programari

Introducció a les BD

Què hi ha en aquest material

2

- Persistència i bases de dades
- Model Lògic relacional (MLR)
- Bases de dades relacionals
 - Esquema relacional
- Anàlisi de les anomalies
 - L'objectiu és evitar anomalies, no pas redundàncies
- Formes normals
 - Dependències funcionals
- Transformació de MC al MLR

Persistència i bases de dades

3

Necessitats de persistència

4

- Gran part de la informació manipulada en el disseny volem que es mantingui fins i tot en el cas d'apagar l'ordinador
 - Diem que volem que la informació sigui **persistent**
 - La persistència s'aconsegueix emmagatzemant la informació en dispositius d'emmagatzematge externs
 - ✦ Disc
 - ✦ Memòria d'estat sòlid
 - ✦ ...

Objectius de les bases de dades

5

- Volem emmagatzemar la informació persistent de tal manera que:

- Sigui útil per a la l'aplicació que l'ha generada
- Pugui ser emprada per aplicacions futures

Independència
de dades

- Exemple

- En un moment determinat hem creat un sistema que ens permet definir caminades i la seva estructuració en etapes
 - ✦ El volem per poder publicar un catàleg on-line amb les nostres propostes
- Més endavant construïm un sistema per poder gestionar les inscripcions a les caminades
 - ✦ Evidentment volem poder usar les dades que ens proporciona el sistema de creació de caminades

Bases de dades

6

- Conjunt d'informació estructurada, persistent, que:
 1. Independitza les dades de les aplicacions
 2. Permet fer els accessos previstos de manera molt eficient
 3. També permet fer altres tipus d'accés

Propietats

7

- Totalitat

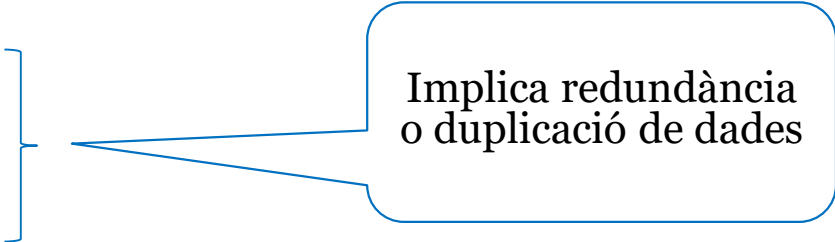
- Conté tota la informació necessària del domini (empresa, negoci, problema,...)

- Consistència

- En tot moment les dades contingudes mantenen les propietats especificades
 - ✦ Cal evitar **anomalies**

- Altres

- Eficiència en espai
- Eficiència en temps
- Robustesa



Implica redundància
o duplicació de dades

Sistema Gestor de Bases de Dades

8

- SGBD, DBMS
- Programari que ens permet crear, gestionar i consultar bases de dades

Base de dades Relacional

9

- Model de bases de dades definit els anys 70
- Propietats bàsiques:
 - Representació gràfica intuïtiva
 - ✦ Permet fer raonaments intuïtius
 - ✦ Permet la fàcil comunicació amb l'usuari no informàtic
 - Model matemàtic
 - ✦ Permet definir una semàntica clara i unívoca de les dades
 - ✦ Permet demostrar la correctesa dels motors dels SGBDR
 - Simplicitat
 - ✦ Permet la construcció de SGBDR comercials
 - ✦ Permet l'ús extens de les BDR

Altres models
de l'època:
jeràrquic i xarxa

Model relacional



Elements estructurals del model relacional

- **Relació**
 - Element bàsic
 - ✦ Es pot expressar en forma tabular
 - ✦ Es pot implementat amb un fitxer
- **Restriccions d'integritat**
 - Hi ha restriccions imposades pel model relacional
 - Hi ha restriccions imposades per la semàntica del problema
 - ✦ Les dependències funcionals
- **Un llenguatge de definició i manipulació**
 - Àlgebra i càlcul relacional
 - ✦ SQL, QBE

Elements dinàmics del model relacional

12

- Mecanismes de comprovació de les restriccions
 - Disparadors
 - Procediments emmagatzemats
- Gestor de transaccions i concurrència
 - Atomicitat, Consistència, Independència, Durabilitat (ACID)
- Gestor de recuperacions
 - Bitàcola, punts de restauració
- Gestor de seguretat
 - Autorització, autenticació

Més que propietats
del model relacional
són eines del
SGBDR

Relació

13

Relació

14

- Donats els conjunts D_1, \dots, D_n (anomenats **dominis**) anomenem **relació** R a tot conjunt de n -tuples **ordenats** tals que cada element del tuple pertany al domini corresponent

$$R = \{ \langle d_1, \dots, d_n \rangle \mid d_i \in D_i \ \forall i \leq n \}$$

$$R \subseteq D_1 \times \dots \times D_n$$

No s'ha de
confondre Relació
amb Interrelació

Concepte matemàtic

Concepte no formal.
Expressa una semàntica del problema. Ex:
associació

Relació: una o moltes?

15

- Sigui la següent definició d'una relació

$R \subseteq \#Prod \times Nom \times Color \times Pes \times Magatzem$

- Són molts els conjunts de tuples que podem expressar amb aquesta definició
 - Amb una definició capturem moltes relacions diferents
 - ✦ Totes les relacions comparteixen els mateixos dominis, i en les mateixes posicions
 - Comparteixen el patró
 - ✦ El que varia és els tuples concrets presents en cada relació
 - S'aplica el patró de diferents maneres

#Prod	Nom	Color	Pes	Magatzem
P1	Cadira	Blau	1200	Vilafranca
P2	Taula	Negre	2000	Cubelles
P3	Llum	Argent	400	Vilafranca
P4	Telèfon	Negre	250	Vilanova

#Prod	Nom	Color	Pes	Magatzem
P1	Llum	Blau	700	Vilafranca
P2	Taula	Negre	2000	Cubelles
P3	Llum	Argent	400	Vilafranca
P4	Telèfon	Negre	250	Vilanova
P5	Taula	Pi	2100	Cubelles



Dues realitzacions d'una mateixa abstracció

Relació: abstracció i realització

16

- Relació com a abstracció

- Cal indicar els dominis i les seves posicions

- ✦ $R_{\text{abs}} = \langle D_1, \dots, D_n \rangle$

- En el nivell de les abstraccions expressem els dominis, i per tant hi ha una igualtat.
- En el nivell de les realitzacions expressem els valors possibles, i per tant hi ha \subseteq

- Relació com a realització

- Conjunt de tuples, cadascun dels quals es pot veure com una realització de la mateixa relació vista com a abstracció

- ✦ $R_{\text{real}} \subseteq \{ \langle d_1, \dots, d_n \rangle \mid d_i \in D_i \ \forall i \leq n \}$

Els D_i s'obtenen de la R_{abs}

Nomenclatura dels nivells d'abstracció

17

- Relació com a **abstracció**: $R = \langle D_1, \dots, D_n \rangle$
 - Relació
 - Variable de relació
 - Esquema de relació
- Relació com a **realització** : $r(R) = \{ \langle d_1, \dots, d_n \rangle \}$
 - Relació
 - Realització
 - Estat de la relació
 - Taula

Notació

- Estat
 - $r(R) \subseteq \text{dom}(d_1) \times \dots \times \text{dom}(d_n) = D_1 \times \dots \times D_n$
- Tuple
 - $t(r) \in r(R)$

Dominis i atributs

18

- Cada **domini** expressa el conjunt de valors que pot tenir un tuple en aquesta posició
- No unicitat dels dominis
 - Diferents posicions poden tenir el mateix domini
- Dominis batejats
 - Sovint posem nom a les posicions del tuple. En aquest cas no es poden repetir noms dins d'una relació:

$$R = \langle a_1:D_1, \dots, a_n:D_n \rangle$$

- **Atribut:**
 - Cadascuna de les posicions dels tuples d'una relació
 - Ús que d'un domini es fa dins d'una relació
 - La semàntica d'un atribut és el rol que assignem al domini en aquesta posició del tuple
 - ✦ És la interpretació dels valors del domini en aquesta posició del tuple

Nomenclatura

19

- **Taula**

- Representació de la realització d'una relació

- **Grau d'una relació**

- Nombre d'elements de cada tuple
- Nombre de columnes de la taula

- **Cardinalitat**

- Nombre de tuples en una realització d'una relació
- Nombre de files de la taula

- **R = Productes =**

{ #Prod, Nom, Color, Pes, Magatzem }

- Cada posició té definit un domini
- Hi pot haver posicions amb el mateix domini

#Prod	Nom	Color	Pes	Magatzem
P1	Cadira	Blau	1200	Vilafranca
P2	Taula	Negre	2000	Cubelles
P3	Llum	Argent	400	Vilafranca
P4	Telèfon	Negre	250	Vilanova

Productes

Qüestions d'ordre

20

- En la realització d'una realització els tuples no estan ordenats
 - L'estat o realització d'una relació és un conjunt de tuples
- Els valors dins d'un tuple estan ordenats
 - Cadascun ha de pertànyer al domini pertinent, segons l'esquema o definició de la relació

Predicat d'una relació

21

- En el modelat, una relació té una semàntica, el **predicat de la relació**
 - Els atributs són les variables lliures del predicat
 - Tot tuple d'una realització de la relació ha de fer cert el predicat de la relació: és una **proposició** (o **enunciat**) certa sobre aquest predicat
 - Una relació és una **asserció** sobre el problema
 - ✦ Cada tuple és un **fet** que compleix aquesta asserció

Exemple

- $R = \text{Producte}(\#Prod, \text{Nom}, \text{Color}, \text{Pes}, \text{Magatzem})$
- **Asserció:** El producte amb codi #prod s'anomena Nom, és de color Color, pesa Pes i el magatzem on es troba és Magatzem
- **Fets:** El producte amb codi P1 s'anomena cadira, és de color blau, pesa 1200gr i es troba al magatzem de Vilafranca

#Prod	Nom	Color	Pes	Magatzem
P1	Cadira	Blau	1200	Vilafranca
P2	Taula	Negre	2000	Cubelles
P3	Llum	Argent	400	Vilafranca
P4	Telèfon	Negre	250	Vilanova

Restriccions d'integritat

22

Súper-clau

23

- Sigui una relació $R = A_1 \times \dots \times A_n$
- Sigui un estat $r(R)$
- Sigui A un subconjunt dels atributs de R :
 $A \subseteq \{A_1, \dots, A_n\}$

Cada atribut A_i és del domini $\text{dom}(A_i)$

- Sigui t_1 i t_2 dos tuples diferents qualssevol de $r(R)$

- Direm que A és una **súper-clau** de la relació R si i $t_1[A] \neq t_2[A]$ per a tot estat $r(R)$

- No hi ha dos tuples diferents que tinguin els mateixos valors en els atributs que conformen la súper-clau
- $t_1[A] = t_2[A] \Rightarrow t_1 = t_2$

Notació

- $t[A]$ és la **projecció** de t sobre A
 - És a dir, el tuple t restringit als atributs indicats a A

Súper-clau (1)

24

- Cap de les tres columnes és una súper-clau

U	V	W
x	a	1
x	b	2
y	a	1
x	b	1

	U	V	W
1	x	a	1
2	x	b	2
3	y	a	1
4	x	b	1

Si prenem una columna qualsevol, hi ha files repetides

Súper-clau (2)

25

- Cap conjunt de dos atributs tampoc és una súper-clau
 - Tenim files repetides

	U	V	W
1	x	a	1
2	x	b	2
3	y	a	1
4	x	b	1

U	V
x	a
x	b
y	a
x	b

V	W
a	1
b	2
a	1
b	1

U	W
x	1
x	2
y	1
x	1

Súper-clau enganyosa

26

- Siguin els conjunts $X=\{U,V\}$ i $Y=\{U,W\}$
 - Són súper-clau?

	U	V	W
1	x	a	1
2	x	b	2
3	y	a	1
4	z	b	1

U	V
x	a
x	b
y	a
z	b

U	W
x	1
x	2
y	1
z	1

No hi ha files repetides

En aquest estat de relació no hi ha files repetides, però per poder afirmar que és una súper-clau hem d'assegurar que **mai**, en cap estat possible, hi haurà files repetides

Detecció de les súper-clau

27

- L'anàlisi d'un estat de relació permet **bandejar** candidats a súper-clau

- Per determinar les **súper-clau** l'única manera és a partir del **predicat de relació**

Clau

28

- Anomenem **clau** de R tota súper-clau de R que és **mínima**
 - La supressió de qualsevol atribut de la clau fa que aquesta deixi de ser súperclau
- Propietats d'una clau
 - **Unicitat**
 - ✦ En tot estat vàlid de la relació i un valor qualsevol de la clau, només hi ha un tuple amb aquest valor
 - **Irreductibilitat**
 - ✦ Cap subconjunt de la clau és clau. La clau és un conjunt mínim d'atributs
- Donada una relació R hi poden haver múltiples claus
 - Totes les claus són iguals
 - ✦ No hi ha unes claus més importants que d'altres

Clau forana

29

- Un conjunt CF d'atributs de R1 és una **clau forana** sobre R2 sii
 - **Compatibilitat**
 - ✦ R2 té una clau que és un conjunt C d'atributs tal que $|CF| = |C|$
 - **Consistència**
 - ✦ Els dominis de CF són els mateixos que els de C
 - Es pot establir una correspondència 1-1 entre els atributs de C i els de CF de tal manera que $f(c_i) = cf_j \Rightarrow \text{dom}(c_i) = \text{dom}(cf_j)$

Restriccions d'integritat

30

- Restriccions de domini

- Condicions sobre el conjunt de valors del domini que es permeten per cada atribut
 - ✦ Ex. Atomicitat (1FN)
 - ✦ Ex. Cada valor és vàlid dins del domini on apareix

El nul com a no valor porta problemes en la unicitat i irreductibilitat de la clau

- Unicitat de les claus

- Integritat d'entitats

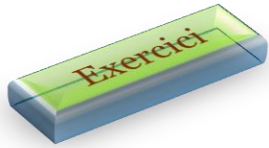
- Les claus no poden ni ser nul·les ni contenir atributs amb el valor nul

- Integritat referencial

- Tota referència que des d'un tuple $t_1(R_1)$ es faci a un tuple $t_2(R_2)$ s'ha de referir a un tuple t_2 existent en l'estat actual $r(R_2)$
 - ✦ Les claus foranes fan referència a un tuple existent, o bé són nul·les

- Restriccions del problema

- Restriccions definides per l'usuari
 - ✦ Ex. El sou no pot ser inferior al sou mínim



Violació de les restriccions d'integritat

31

- Construiu dues relacions, una amb una clau forana cap a l'altra
- Definiu operacions d'**inserció** que violin:
 - Unicitat de les claus
 - Restricció d'entitat
 - Restricció de integritat referencial
- Definiu operacions d'**esborrat** que violin:
 - Restricció de integritat referencial
- Definiu operacions d'**actualització** que violin:
 - Unicitat de les claus
 - Restricció d'entitat
 - Restricció de integritat referencial

Base de dades relacional

32

Base de dades

33

- Esquema de relació:

- Nom de la relació +

- Conjunt d'atributs

- ✦ Cada atribut és un parell $\langle \text{nom}, \text{domini} \rangle$

Esquema de relació = Relació com a abstracció

- Esquema de **base de dades**:

- Conjunt S d'esquemes de relacions +

- Conjunt de restriccions d'integritat sobre S

- El SGBDR s'ha de preocupar que qualsevol modificació de l'estat actual de la BD mantingui les restriccions d'integritat

- La BD sempre ha d'estar en un estat íntegre o consistent

Exemple d'esquema de BD

34

Client = <#Client, NomClient, CiutatClient>



Les claus s'indiquen
amb subratllat

Venda = <#Client, #Prod, quantitat>



Producte = <#Prod, NomProd, Color, Pes, CiutatMagatzem>

No considerem cap més restricció d'integritat que les restriccions de domini, d'entitat i d'integritat referencial

Model lògic relacional

35

- Anomenem **model lògic relacional** a l'esquema de base de dades en termes de relació

Anomalies

36

Anomalia de BD

37

- **Anomalia:** Situació en què el disseny de la Base de Dades dificulta o impedeix de mantenir correctament la informació
 - ✦ Dificultat o impossibilitat de mantenir la **semàntica** de les dades
 - ✦ Dificultat o impossibilitat de mantenir les **restriccions d'integritat**
 - ✦ Dificultat o impossibilitat de mantenir la **informació** coneguda
- L'estudi de les anomalies permet cercar millors dissenys per a la BD

Tipus d'anomalies

38

- Treballador(#Treb, nomTreb, dataContractació, #Departament, adreçaDep)
- **Inserció**
 - Problemes per assegurar la consistència
 - ✦ En totes les contractacions d'un mateix departament hem d'introduir correctament l'adreça del departament,
 - Això cal per mantenir consistent la semàntica de les dades
 - Introducció de nuls
 - ✦ Si és possible la contractació sense assignació immediata a cap departament, caldrà introduir nuls
 - Els nuls dificulten el manteniment de les restriccions d'integritat
 - Violació de la restricció d'entitat
 - ✦ La creació d'un nou departament exigeix introduir un tuple que conté nuls a la clau
 - Impossibilitat de mantenir una restricció d'integritat
- **Esborrat**
 - Pèrdua indesitjada d'informació
 - ✦ Si eliminem l'únic treballador d'un departament també estem eliminant la informació sobre l'adreça del departament
 - Impossibilitat o dificultat de mantenir la informació
- **Actualització**
 - Exigència de propagar els canvis
 - ✦ L'actualització o modificació del nom d'un departament en un treballador significa modificar el nom del departament en tots els treballadors d'aquest departament
 - Això cal per mantenir consistent la semàntica de les dades

El perquè de les anomalies

39

- La relació que provoca l'anomalia té un **predicat** que de fet és la conjunció de predicats més simples
 - Cada fet o enunciat és per tant la conjunció d'altres fets més simples, que anomenarem constituents
 - Exemple
 - ✦ Treballador(...) = En Pere és treballador des de tal dia, i treballa en el departament 4, el de Recerca i innovació
 - ✦ Treballador(...) =
Empleat(...) \wedge Departament(...) \wedge Assignació(...)
 - Empleat() = En Pere és un treballador des de tal dia
 - Departament = El departament 4 és el de Recerca i Innovació
 - Assignació(...) = En Pere està assignat al departament 4
- Solució
 - Descompondre la relació “anòmala” en d'altres relacions
 - ✦ Es tracta de passar d'un predicat compost a un conjunt de predicats atòmics

Exemple: Anomalia d'inserció

- Volem inserir un enunciat sobre un conjunt no total dels constituents
 - Assignació()
 - Hem de modificar un tuple ja existent
 - Empleat()
 - Hem de deixar a nul la informació del departament
 - Departament()
 - Hem de deixar a nul la informació de l'empleat

Anomalies des del punt de vista de la informació

40

- **Pèrdua d'informació**

- Cas particular d'anomalia d'esborrat

- ✦ Volem negar un fet, i només ho podem fer negant un fet complex que el conté

- **Redundància**

- Cas particular d'anomalia d'inserció o d'actualització

- ✦ Volem establir un fet (nou o bé modificar un fet anterior) i només ho podem fer establint un fet complex que el conté

- El fet que ens veiem obligats a establir no pot contenir falsedats. Per això repetim un fet ja conegut. Exemple:

- $P(x,y) = Q(x) \wedge R(y)$
- Tenim un fet sobre P que ens assegura R(a). Per exemple tenim un tuple t_1 a P tal que $t_1 = \langle c, a \rangle$, que assegura $Q(c) \wedge R(a)$
- Volem establir Q(b)
- Com que només ho podem fer a través de tuples de P hem de construir un fet $Q(b) \wedge R(y)$ on R(y) sigui cert segur. Ho fem usant R(a)
- El resultat és que inserim un tuple t_2 a P tal que $t_2 = \langle b, a \rangle$, que assegura el fet $Q(b) \wedge R(a)$
- Per tant R(a) és redundant: l'afirmem tant a t_1 com a t_2

Redundància?

41

- Cal dissenyar una base de dades amb el mínim possible d'anomalies
 - L'objectiu es reduir les anomalies
 - ✦ Com a conseqüència reduïm les redundàncies
 - De fet, només un tipus de redundància
- Sovint es diu que les BDR són un model que pretenen evitar les redundàncies
 - Des del nostre punt de vista això és incorrecte: la bitàcola o els punts de restauració són casos de redundància extrema, desitjada i volguda. Com també ho són els índexs
 - El fonamental és evitar les anomalies
 - ✦ En fer-ho reduïm redundàncies, però no les evitem; tampoc ho pretenem

Formes normals

42

Formes normals

43

- Per tal d'evitar les anomalies **el Model Lògic Relacional** ha de satisfer unes determinades propietats
 - Exigim que cada relació del model estigui en el que anomenem una forma normal
 - La normalització és el que distingeix el model conceptual del lògic

Formes normals

44

- En Codd proposà un procés de transformació en tres etapes:
 - 1FN
 - 2FN
 - 3FN
- Aviat es veié un error en la seva formulació, que obligà a introduir una nova FN, la de Boyce-Codd
 - BCNF
- Les formes normals 3FN i BCNF es poden definir, i obtenir, sense necessitat de les formes 1FN i 2FN

Condicions sobre els dominis

45

- Sobre els dominis dels diferents atributs d'una relació d'un esquema relacional exigim:
 - No estructurats o atòmics
 - ✦ Un valor no pot ser un tuple
 - Monoavaluats
 - ✦ Un valor no pot ser un conjunt
 - Aquesta exigència pot ser discutible i controvertida, però és la base del Model Relacional
- Donada una relació R, si tots els atributs són **atòmics** i **monoavaluats** direm que R està en **1FN** (primera forma normal)

Normalització

46

- S'anomena **normalització** el procés de convertir un esquema relacional en un altre esquema on totes les relacions estiguin en forma normal
 - La 3FN sempre és possible
 - La BCFN sempre és possible, però en suprimir determinades anomalies en pot introduir d'altres

Altres formes normals

47

- Per resoldre determinades anomalies s'han proposat d'altres formes normals, basades en generalitzacions de les dependències funcionals
 - 4FN
 - 5FN

Pros i contres de la normalització

48

- La normalització és un procés feixuc i complex, tot i que mecanitzable
- Per poder normalitzar cal conèixer totes les restriccions que la semàntica del problema imposa sobre les dades
 - En concret ens preocupem de les restriccions expressable en forma de **Dependència Funcional**
 - ✦ Les claus són un tipus de dependència funcional
 - Cal per tant una tasca manual prèvia
- L'esquema relacional resultant pot contenir relacions de difícil comprensió semàntica
 - Es fa difícil definir-ne el predicat semàntic
 - Per a l'usuari pot ser complex entendre el paper de la relació

Dependències funcionals

49

Dependència funcional

50

- Sigui R una **relació**
 - ✦ De fet un esquema de relació
- Sigui $\text{atr}(R)$ el conjunt de tots els atributs de R
- Siguin $A \subseteq \text{atr}(R)$ i $B \subseteq \text{atr}(R)$
- Sigui $r(R)$ un **estat** qualsevol de R
- Siguin t_1 i t_2 dos **tuples** qualssevol de $r(R)$
- Direm que existeix una **dependència funcional** $A \rightarrow B$ si:

$$t_1[A]=t_2[A] \Rightarrow t_1[B] = t_2[B]$$

- En un estat qualsevol de la relació no hi pot haver dos tuples diferents que coincideixin en els atributs de A però no en els de B
 - ✦ La coincidència en els valors de A significa la coincidència també en els valors de B

Dependència funcional i súper-clau

51

- Si $K \subseteq \text{atr}(R)$ és una súper clau de R , llavors $K \rightarrow \text{atr}(R)$
 - Demostració
 - ✦ Per definició de súper-clau, $t_1[K]=t_2[K] \Rightarrow t_1=t_2$
 - ✦ Per tant, $t_1[K]=t_2[K] \Rightarrow t_1[\text{atr}(R)]=t_2[\text{atr}(R)]$

Tot tuple depèn funcionalment de les seves súper-claus

- La definició de súper-clau és un cas particular de dependència funcional

El paper de les dependències funcionals

52

- Imposen restriccions als fets que es poden establir
 - Imposen restriccions sobre els tuples que podem tenir a la base de dades
- **Són restriccions semàntiques**
 - L'usuari, com a coneixedor de les dades, és qui les pot definir

Les dependències funcionals són **restriccions d'integritat** que **exigim** a la BD

$sou > 4$ és una restricció d'integritat que no és expressable en termes de dependència funcional

Transformació del MC al model relacional

53

Objectiu de la transformació

54

- Tenim un model conceptual, MC
- Volem construir una BD relacional capaç de mantenir aquest MC
 - Necessitem un conjunt de relacions normalitzades
 1. Transformació Conceptual → Lògic
 2. Normalització
- Si partim d'un MC ben construït el model lògic resultant de la transformació està (gairebé) normalitzat
 - Cal una comprovació final, que ens pot permetre detectar errades

Idea bàsica

55

- **Cada element E del model conceptual es transforma en una relació R**

- **Avantatges**
 - Ortogonalitat, simplicitat
 - Es manté la semàntica de MC

Context

56

- Valor de la funció `Context` segons cada cas
 - Concepte A: \emptyset
 - Associació entre A i B: $\{A, B\}$
 - Concepte associatiu entre A i B: $\{A, B\}$
 - Especialització B de A: $\{A\}$
 - Concepte feble B de A: $\{A\}$

ClauInduïda

57

- Valor de la funció ClauInduïda
 - Concepte
 - ✦ $\text{ClauInduïda}(A) = \text{Una}$ de les claus de la relació RA en què s'ha transformat A
 - Conjunt de conceptes: $A = \cup A_i$
 - ✦ $\text{ClauInduïda}(A) = \cup \{ \text{ClauInduïda}(A_i) \}$

La clau induïda d'un conjunt de conceptes és un **conjunt de claus**, una per cadascun dels conceptes

Concatenació d'un conjunt

58

$$\text{concat}(\{A_i\}) = \cup A_i$$

- La concatenació d'un conjunt de conjunts és el conjunt que conté junts els elements de tots els conjunts concatenats
 - La concatenació el que fa és aplanar el conjunt
 - La concatenació d'un conjunt de claus (on una clau és un conjunt d'atributs) és un sol conjunt d'atributs
- Exemple
 - $\text{concat}(\{\{a,b,c\}, \{a,d\}, \{e,f\}\}) = \{a,b,c,d,e,f\}$

Transformació MC→Relacional

59

Cada element E de MC es transforma en una relació R

- $\text{atr}(R) = \text{atr}(E) \cup \text{conc}(\text{ClauInduïda}(\text{context}(E)))$
- $\text{Clau}(R) = \text{conc}(\text{filtre}(\text{FontClau}(R)))$

- **Atributs de la relació**
 - Els propis més els que formen les claus induïdes pel context
 - ✦ Prenem una clau per cada element del context
- **Claus de la relació**
 - Ho analitzem a continuació

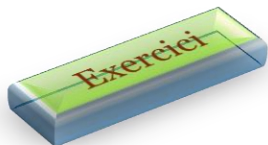
Font de les claus

60

$$\text{FontClau}(R) = \text{ClauInduïda}(\text{context}(E)) \cup \{\{\text{id}(E)\}\}$$

- Clau = conjunt atributs
- Clau induïda del context = conjunt de claus
 - Una clau per cada element del context
 - Per tant és un conjunt on cada element és un conjunt d'atributs
- $\text{FontClau}(R)$
 - En el conjunt de claus hi afegim el conjunt d'un sol element $\{\text{id}(E)\}$
 - És un conjunt de conjunt d'atributs

- La font de la clau és un conjunt de claus.
- Les claus d'aquest conjunt són:
 - Un identificador propi
 - Una clau presa de cadascun dels elements del context



Font de la clau

61

- **Suposem**

- $\text{Context}(A) = \{B, C\}$
- $\text{Clau}(B) = \{ \{b_1, b_2\}, \{b_3\} \}$
 - ✦ B té dues claus, una d'un atribut i l'altre de dos
- $\text{Clau}(C) = \{ \{c_1\} \}$
 - ✦ c té una sola clau, formada per un sol atribut

- **Llavors**

- $\text{ClauInduïda}(\text{context}(A)) =$
 $\text{ClauInduïda}(B) \cup \text{ClauInduïda}(C) =$
 $\{ \{b_1, b_2\} \} \cup \{ \{c_1\} \} = \{ \{b_1, b_2\}, \{c_1\} \}$
- $\text{FontClau}(A) =$
 $\{ \{b_1, b_2\}, \{c_1\} \} \cup \{ \text{id}_A \} =$
 $\{ \{b_1, b_2\}, \{c_1\}, \{ \text{id}_A \} \}$

Prenem només una de les dues claus

Filtre

62

- El filtre el que fa és esborrar de la font de la clau aquelles claus innecessàries
 - De totes les claus induïdes per elements del context que participen a E amb multiplicitat 1, n'esborrem una qualsevol
 - Si el context no és buit i E no és un concepte feble, esborrem $\{idE\}$
 - ✦ Mantenim $\{IdE\}$ en qualsevol dels següents casos
 - Context buit
 - E és un concepte
 - Concepte feble
 - El context no és suficient



Selecció

63

- Suposem un concepte associatiu A' entre B' i C'

- **Llavors**

- $\text{Context}(A) = \{B, C\}$
- $\text{Clau}(B) = \{\{b_1, b_2\}, \{b_3\}\}$
 - ✦ B té dues claus, una d'un atribut i l'altre de dos
- $\text{Clau}(C) = \{\{c_1\}\}$
 - ✦ c té una sola clau, formada per un sol atribut

- X' expressa un element del MC
- X expressa la relació del ML en la que es transforma X'

- **Sabem que**

- $\text{ClauInduïda}(\text{context}(A)) = \{\{b_1, b_2\}\} \cup \{\{c_1\}\} = \{\{b_1, b_2\}, \{c_1\}\}$
- $\text{FontClau}(A) = \{\{b_1, b_2\}, \{c_1\}\} \cup \{\text{idA}'\} = \{\{b_1, b_2\}, \{c_1\}, \{\text{idA}'\}\}$

- **Cas M-N**

- El context és fort. Esborrem $\{\text{idA}'\}$
- $\text{Clau}(A) = \text{concat}(\{\{b_1, b_2\}, \{c_1\}\}) = \{b_1, b_2, c_1\}$

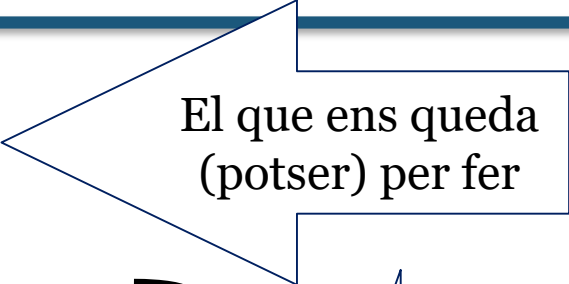
- **Cas N-1**

- El context és fort.: esborrem $\{\text{idA}'\}$
- C té multiplicitat 1: esborrem $\{c_1\}$
- $\text{Clau}(A) = \text{concat}(\{\{b_1, b_2\}\}) = \{b_1, b_2\}$

Pocs graus de llibertat

64

- Concepte, entitat, “classe” → Relació
 - Les claus venen determinades per la semàntica del concepte
- Associació, interrelació → Relació
 - Les claus venen determinades per la informació del MC
- Concepte feble → Relació
 - Les claus venen determinades per la informació del MC
- Concepte associatiu → Relació
 - Les claus venen determinades per la informació del MC
- Especialització (Generalització) → Relació
 - Les claus venen determinades per la informació del MC

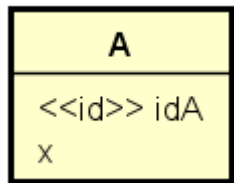


El que ens queda
(potser) per fer

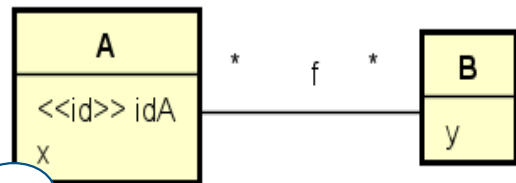


Feina feta

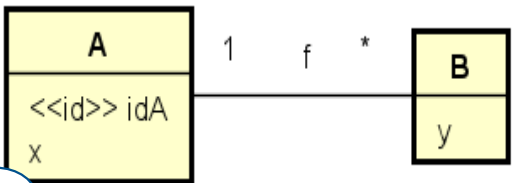
Transformacions bàsiques (1)



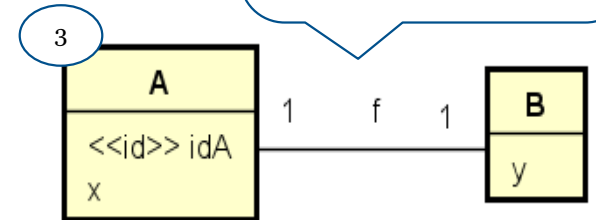
$$R_A = \{\underline{id}_A, x\}$$



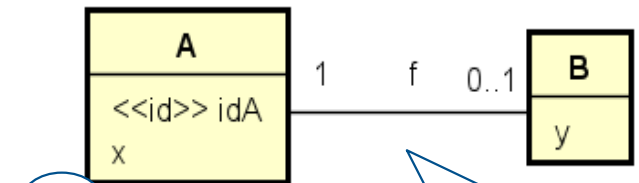
$$\begin{aligned} R_A &= \{\underline{id}_A, x\} \\ R_B &= \{\underline{id}_B, y\} \\ R_f &= \{\underline{id}_A, \underline{id}_B\} \end{aligned}$$



$$\begin{aligned} R_A &= \{\underline{id}_A, x\} \\ R_B &= \{\underline{id}_B, y\} \\ R_f &= \{\underline{id}_A, \underline{id}_B\} \end{aligned}$$



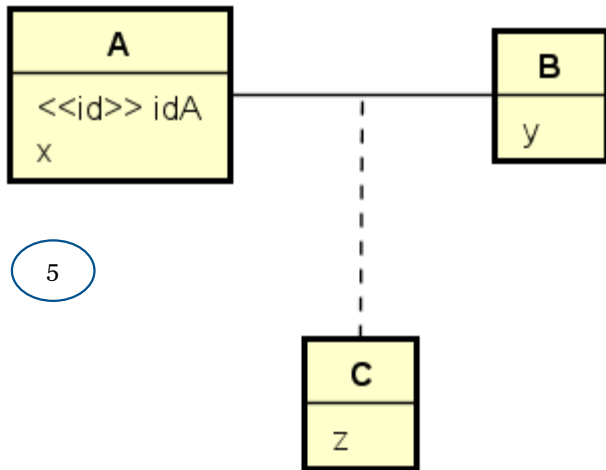
$$\begin{aligned} R_A &= \{\underline{id}_A, x\} \\ R_B &= \{\underline{id}_B, y\} \\ R_f &= \{\underline{id}_A, \underline{id}_B\} \end{aligned}$$



$$\begin{aligned} R_A &= \{\underline{id}_A, x\} \\ R_B &= \{\underline{id}_B, y\} \\ R_f &= \{\underline{id}_A, \underline{id}_B\} \end{aligned}$$

Transformacions bàsiques (2)

69



$$R_A = \{\underline{id}_A, x\}$$

$$R_B = \{\underline{id}_B, y\}$$

$$R_C = \{\underline{id}_A, \underline{id}_B, z\}$$

6



$$R_A = \{\underline{id}_A, x\}$$

$$R_B = \{\underline{id}_A, \underline{id}_B, y\}$$

Lectura de les transformacions bàsiques

70

1. Una única clau, formada per dos atributs
2. Una clau
 - La monoavaluació exigeix suprimir una de les claus induïdes
3. Dues claus
 - Hi ha dues maneres de suprimir una de les claus induïdes sobre una pota a 1
4. Una clau
 - L'optativitat no afecta per a res quina és la clau resultant
5. Una clau
 - La clau només està formada per les claus induïdes del context
6. Dues claus
 - L'especialització tant pot identificar-se dins l'especialització, com dins la generalització

Variant associació 1-N

71

- Transformació canònica

- RA(X,Y)
- RB(Z,T)
- RAB(X,Z,U)

La U tracta el cas d'inrerrelacions amb atributs (els conceptes associatius en terminologia UML)

- Variant

- RA(X,Y,Z,U)
- RB(Z,T)

- Condicions

- La pota B de l'associació compleix que:

- ✦ Té multiplicitat 1
- ✦ És obligatòria

Per assegurar la restricció d'unicitat de la clau

Per evitar la presència de nuls

- Avantatges de la variant

- S'eviten moltes combinacions de taules en les consultes

- Inconvenients de la variant

- La semàntica de RA deixa de ser clara
- Dificultat de promocionar a M-N
- Dos esquemes de transformació

Recomanacions per a la normalització

74

- Per a construir el ML usem la transformació canònica
 - Cada element de MC esdevé una relació a ML
- En etapes molt tardanes del desenvolupament ens plantegem qüestions d'eficiència
 - Introduïm les variants de les transformacions per millorar l'eficiència
 - No és una necessitat lògica, més aviat física

Nivells d'una BD

75

- **Nivell conceptual**
 - Conjunt de dades i interrelacions, expressats des del punt de vista de l'usuari
- **Esquema extern**
 - Fragment del nivell conceptual adaptat a un tipus determinat d'usuari
- **Nivell lògic**
 - Definició de les relacions i dependències funcionals que configuren la BD
- **Nivell físic**
 - Descripció dels fitxers necessaris per a mantenir la informació

Pautes de cara al disseny d'una BDR

76

- **Semàntica propera al problema**
 - El predicat de cada relació ha de ser simple d'exposar
 - ✦ Modificabilitat, mantenibilitat
- **Predicats atòmics**
 - Hem de dissenyar la BDR de tal manera que no hi hagi anomalies d'inserció, esborrat o modificació.
 - ✦ Cada relació té un predicat semànticament atòmic
 - L'ús no genera violacions de les restriccions d'integritat
 - Una operació simple no desencadena múltiples operacions internes
- **Obligatorietat de valor en els atributs**
 - En el predicat de relació tots els arguments (atributs) són rellevants per a tots els tuples
 - ✦ Evitem, tant com sigui possible, els atributs que admeten nuls
 - La semàntica dels nuls és diversa, i gens clara
 - Els nuls porten força problemes