Algebra Relazionale 3

Università degli Studi del Sannio

Facoltà di Ingegneria

Corso di Laurea in Ingegneria Informatica

Corso di Basi di Dati

Anno Accademico 2005/2006

docente: ing. Corrado Aaron Visaggio

email: visaggio@unisannio.it

ricevimento: mercoledì 11.00-13.00.

Relazione Inversa

Assegnati due insiemi A e B diremo:

- **relazione (binaria)**, su A e B: *un sottoinsieme del prodotto cartesiano* $A \times B$, cioè $R \subseteq A \times B$
- (notazioni alternative **(a,b)∈R** oppure **a R b** con $a \in A$ $b \in B$)

gli insiemi A e B sono i **domini della relazione.**

Data una relazione $\mathbf{R} \subseteq \mathbf{A} \times \mathbf{B}$ diremo

- **relazione inversa** di R: la relazione $R^{-1} \equiv \{ (b,a) : (a,b) \in R \}$

Relazione raggiungibile

Data una relazione R $\subseteq$ A×B, altre proprietà sono le seguenti:

- R **simmetrica** $\Leftrightarrow \forall$ *(a,b) ∈ R → (b,a) ∈ R*
- R è **transitiva** $\Leftrightarrow \forall$ *(a,b), (b,c) ∈ R → (a,c) ∈ R*
- **b∈B** è **raggiungibile**, secondo R, da **a∈A**: *se esistono in B n elementi tali che* $(a,b_1)\in R$, $(b_1,b_2)\in R$, $(b_2,b_3)\in R$, …. $(b_n,b)\in R$
- **chiusura transitiva di R**: *l'insieme* $R^+ \equiv \{(a,b)$ *: b è raggiungibile da a}*

la chiusura transitiva di R è una relazione transitiva, mentre non è detto che R lo sia.

ad esempio se A è l'insieme dei padri e B quello dei figli, la relazione R ≡ { (a,b) : a è padre di b} è non transitiva mentre R^+ ≡ { (a,b) : b è discendente di a} lo è.

Chiave / Superchiave

chiave di R: *un sottoinsieme S degli attributi dello schema di R che goda delle proprietà di*:

- **univocità**: *nessuna istanza di R può possedere due tuple che abbiano gli stessi valori in tutti gli attributi di S ;*
- **minimalità**: *non esiste un sottoinsieme degli attributi di S che goda della proprietà di univocità.*

superchiave di R: *un insieme di attributi S che goda della proprietà di univocità ma non di quella di minimalità*

Chiave Primaria

Poiché in una relazione non ci possono essere due tuple uguali, ogni relazione ha come chiave l'insieme di tutti i suoi attributi.

- **chiave candidata**: *una qualsiasi delle chiavi della relazione*
- **chiave prima**: *la chiave che viene adottata per l'accesso a tuple della relazione*.

Nella pratica una tupla di una relazione può avere, per uno o più attributi, un **valore non specificato** (Φ) allora:

- **chiave primaria**: *una chiave i cui attributi non posseggano in nessuna istanza della relazione ed in nessuna tupla di tale istanza un valore Φ.*
- **attributo primo**: *attributo che fa parte di almeno una chiave primaria.*

Dipendenza Funzionale...

Una chiave prima deve essere anche primaria.

Un attributo **Y dipende funzionalmente da** un attributo **X**:*se Y=F(X)* ovvero *se non esistono istanze della relazione nelle quali si abbia che due tuple che presentano lo stesso valore per X non abbiano anche uguale valore in Y*.

Diremo, equivalentemente, che **X determina funzionalmente Y** ed useremo la notazione **X→Y**. La definizione si estende, in maniera analoga, alla **dipendenza funzionale da un insieme di attributi**.

...Dipendenza Funzionale

La **dipendenza funzionale X→Y** si dice

- **completa**: *se Y non dipende funzionalmente da un sottoinsieme proprio di X;*
- **transitiva**: *se esiste un attributo Z diverso da Y e da X per cui si abbia X→Z and Z→Y.*

Detto F un insieme di dipendenze funzionali dello schema di R, si dice

- **F implica logicamente X→Y** (**F** ⇒ **X→Y**): *se ogni istanza di R che soddisfa le dipendenze in F soddisfa anche X→Y;*
- **chiusura transitiva F^+** di F: *è l'insieme delle dipendenze implicate da F.*

Seconda e Terza Forma Normale

relazione in 2ª f.n.: *se è in 1ª f.n. e se tutti i suoi attributi non primi dipendono funzionalmente e completamente dalla chiave;*

BDR in 2ª f.n.: *se lo è ogni sua relazione;*

relazione in 3ª f.n.: *se è in 2ª f.n. e se ogni suo attributo non primo non dipende transitivamente dalla chiave;*

BDR in 3ª f.n.: *se lo è ogni sua relazione.*

Calcolo Relazionale

Il Calcolo Relazionale fa riferimento ad una famiglia di linguaggi di interrogazione **dichiarativi**, basati sul calcolo dei **predicati del primo ordine**.

Nella logica matematica il linguaggio del primo ordine è un linguaggio formalizzato che serve per gestire enunciati e ragionamenti che coinvolgono i connettivi logici, le proprietà e i quantificatori "per ogni ..." () ed "esiste..." ().

L'espressione "**del primo ordine**" indica che c'è un insieme di riferimento e i quantificatori possono riguardare solo gli elementi di tale insieme e non i sottinsiemi.

Ad esempio si può dire "per tutti gli x elementi dell'insieme vale P(x)" ma non si può dire "per tutti i sottoinsiemi A vale P(A)"; quando si permette che i quantificatori possono spaziare anche tra tutti i possibili sottoinsiemi dell'insieme di riferimento si parla di **teoria del secondo ordine**.

L'algebra Relazionale

L'algebra relazionale è un linguaggio procedurale in quanto le sue espressioni specificano attraverso le singole applicazioni degli operatori la costruzione del risultato.

Si presenterà qui il calcolo **relazionale sui domini**, con le seguenti differenze:

- i simboli di predicato corrispondono alle relazioni
- i simboli di funzione non compaiono.

Nel calcolo dei predicati interessano formule aperte (con variabili libere) e formule chiuse (con variabili vincolate).

Nel calcolo relazionale interessano prevalentemente le formule aperte.

Linguaggio del primo ordine

Un linguaggio del primo ordine è caratterizzato da:

- un alfabeto di simboli
- un insieme di termini (che dovrebbero denotare gli "oggetti" dell'insieme che si sta considerando)
- un insieme di formule ben formate (fbf) cioè un insieme di stringhe composte di simboli dell'alfabeto che vengono considerate sintatticamente corrette

L'insieme dei termini è costituito da tutte quelle stringhe dell'alfabeto che denotano degli oggetti specifici:

- una costante individuale è un termine
- una variabile è un termine
- se t1,...,tn sono n termini e f è un simbolo per funzione n-aria allora f(t1,...,tn) è un termine
- nient'altro è un termine

L'Alfabeto

- simboli per variabili (infiniti): x1,x2,x3,...
- simboli per costanti individuali (eventualmente nessuno): a1,a2,...
- simboli per predicati (o relazioni), a ciascuno dei quali è associato il suo numero di argomenti: P1,Q1,P2,Q2,...
- simboli per funzioni, a ciascuno dei quali è associato un numero corrispondente al numero di argomenti: f1,g1,f2,g2,...
- simboli di punteggiatura: "(", ")" e la virgola ","
- simboli per connettivi logici: (negazione), (implicazione), (se e solo se), (e), (oppure)
- simboli per quantificatori: (quantificatore universale), (quantificatore esistenziale)

Calcolo relazionale sui domini...

Le espressioni del calcolo relazionale sui domini hanno la seguente formula:

$$\{A_1 : x_1, \dots, A_k : x_k \mid f\}$$

dove

$A_1, \dots, A_k$ sono attributi distinti (target list)

$x_1, \dots, x_k$ sono variabili (e rendono vera la formula)

f è una formula tale che*:*

- *Vi sono formule atomiche di due tipi:*
  - $R(A_1 : x_1, \dots, A_p : x_p)$, dove $R(A_1, \dots, A_p)$ è uno schema di relazione e $x_1, \dots, x_p$ sono variabili
  - $x \vartheta y$ $x \vartheta c$ $(=, \neq, \leq, \geq, <, >)$ con x e y variabili e c costante

...Calcolo relazionale sui domini

- Se f_1 f_2 sono formule, lo sono anche $f_1 \wedge f_2$ $f_1 \vee f_2$ $\neg f_2$;
- Se f è una formula ed x una variabile, allora anche $\exists x(f)$ $\forall x(f)$ sono formule.

- Una formula
  - $R(A_1 : x_1, \dots, A_p : x_p)$, dove $R(A_1, \dots, A_p)$ è vera sui valori di $x_1, \dots, x_p$ che formano una tupla di R;
  - una formula atomica $x \vartheta y$ $x \vartheta c$ è vera sui valori a_1, a_2 se i confronti $a_1 \vartheta a_2$ e $a_2 \vartheta c$ sono soddisfatti.
  - Per congiunzione, disgiunzione e negazione valgono le usuali definizioni.
  - $\exists x(f)$ è vera se esiste almeno un valore a di x che renda vera f.
  - $\forall x(f)$ è vera se per ogni valore a di x f è vera.

Un Esempio

- Interrogazione: *trovare il titolo dei film, il cognome e nome del regista, il cui film è stato prodotto dopo il 1960.*

$\pi_{\text{Titolo, Cognome Regista, Nome Regista}} (\sigma_{\text{Anno} > 1960}(\text{FILM}))$

{Titolo: t, Cognome Regista: c, Nome Regista: n |

FILM (Titolo: t, Cognome Regista: c, Nome Regista: n, Anno : a) $\wedge$ a >1960}

- tcna costituisca una tupla di FILM
- il valore di a sia maggiore di 1960

Limiti del calcolo relazionale

Il calcolo ammette espressioni che hanno poco senso!

$$\{A_1 : x_1, A_2 : x_2 \mid R(A_1 : x_1) \vee x_2 = x_2\}$$

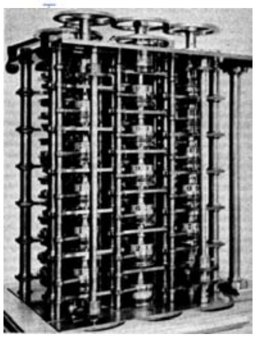
Le espressioni dipendenti dal dominio non hanno utilità pratica: i linguaggi di interrogazione devono essere indipendenti dal dominio.

L'algebra relazionale è indipendente dal dominio perché costruisce il risultato a partire dalla relazione, a differenza del calcolo relazionale.

Il calcolo relazionale presenta troppe variabili, una per ogni attributo coinvolto nell'interrogazione.

Si introduce il calcolo su tuple con dichiarazioni di range, ove le variabli indicano tuple invece di attributi.

Charles Babbage



Matematico, nato a Londra nel 1792(o 91?).

All'età di 24 anni fu eletto membro della Royal Society

Studiava astronomia e elettro-magnetismo.

Inventa la differential engine, che consente di fare operazioni complesse utilizzando solo addizioni.

La macchina calcolava la serie $n^2 + n + 41$.

I numeri sono: 41, 43, 47, 53, 61, ... Le differenze dei termini sono 2, 4, 6, 8, … e le seconde differenze sono 2, 2, 2, ...

Voleva calcolare i logaritmi in modo "meccanico" (non automatico)

Muore nel 1871.