



Consorzio per la formazione e la ricerca in Ingegneria dell'Informazione

## *Introduzione al Modello Relazionale*

Semplicità e potenza

Docente:  
Cesare Colombo  
CEFRIEL  
colombo@cefriel.it  
<http://www.cefriel.it/~colombo>

Basi di Dati



### Origini del modello



- La maggior parte dei sistemi odierni si basa sul modello relazionale
- Il modello relazionale nasce nel '70 su solide basi matematiche
  - ▶ L'estrema astrazione del modello lo rende di fatto poco utilizzato per i primi 10 anni
  - ▶ Si afferma comunque perché è il primo modello che porta ad avere la proprietà di **indipendenza dei dati**



## Concetti fondamentali



---

- Il modello relazionale si basa su due concetti fondamentali:
  - ▶ **Relazione** → concetto matematico, teoria degli insiemi → **Potenza**
  - ▶ **Tabella** → concetto semplice ed intuitivo → **Semplicità**
- I due concetti sono riconducibili l'uno all'altro e questo ne ha decretato il successo: anche gli utenti finali ne comprendono il significato
- **Indipendenza dei dati**: gli utenti e i programmatori vedono solo il livello logico (ciò non era possibile con gli altri modelli)

---

Il modello Relazionale
- 3 -
Basi di Dati



## Relazione matematica



---

- Dati due insiemi D1 e D2, si chiama **prodotto cartesiano**  $D1 \times D2$  l'insieme di tutte le coppie ordinate di valori provenienti da tali insiemi
- Una **relazione matematica** è un sottoinsieme delle coppie

**Esempio di prodotto**

- $A = \{1, 2, 4\}$
- $B = \{a, b\}$
- $A \times B = \{(1, a), (1, b), (2, a), (2, b), (4, a), (4, b)\}$

**Esempi di relazione**

- $R1 = \{(1, a), (4, a)\}$
- $R2 = \{(1, b), (4, a), (4, b)\}$

NOTA: Tutte le definizioni possono essere generalizzate a più insiemi

---

Il modello Relazionale
- 4 -
Basi di Dati



## Tabelle



---

- Le relazioni possono essere rappresentate in modo efficace tramite tabelle
- $R1 = \{(1,a), (4,a)\}$
- $R2 = \{(1,b), (4,a), (4,b)\}$
- Definizioni
  - ▶ **Domini:** insiemi su cui è definita la relazione
  - ▶ **Grado:** numero delle componenti del prodotto (e quindi di ogni n-upla)
  - ▶ **Cardinalità:** numero di elementi della relazione (= numero di n-uple)

1	a
4	a

1	b
4	a
4	b

---

Il modello Relazionale

- 5 -

Basi di Dati



## Esempio



---

- Le relazioni possono essere usate per rappresentare i dati di interesse per qualche applicazione
- Es: Dati relativi ad un insieme di partite di calcio

Inter	Juventus	2	0
Lazio	Roma	0	0
Inter	Milan	6	0
Bari	Milan	0	1

- Si può vedere come un sottoinsieme del prodotto cartesiano

String x String x Intero x Intero

---

Il modello Relazionale

- 6 -

Basi di Dati



## Relazioni con Attributi (1)



- Quando usiamo le relazioni per rappresentare i dati ogni n-upla contiene dati fra loro collegati
- Nella definizione di relazione matematica è importante la nozione di ordine
  - ▶ Inter - Juve → 3 - 0  
è ben diverso da  
Juve - Inter → 3 - 0
- Quando non indispensabile, in informatica si tende a utilizzare notazioni *non posizionali*
  - ▶ E' più semplice e intuitivo
- Si definiscono allora le **relazioni con attributi**



## Relazioni con Attributi (2)



- Si associa a ciascun insieme del prodotto cartesiano un nome, **attributo**, che descrive il ruolo giocato dall'insieme stesso
  - ▶ I nomi devono essere diversi l'uno dall'altro
  - ▶ Ogni riga viene detta "tupla"

Squadra di Casa	Squadra Ospitata	Reti SC	Reti SO
Inter	Juventus	2	0
Lazio	Roma	0	0
Inter	Milan	5	0
Bari	Milan	0	1



## Tabelle e Relazioni



- Una tabella rappresenta una relazione se
  - ▶ i valori di ciascuna colonna sono fra loro omogenei (dallo stesso dominio)
  - ▶ le righe sono diverse fra loro (è un'insieme)
  - ▶ le intestazioni delle colonne sono diverse tra loro (attributi)
- Inoltre, in una tabella che rappresenta una relazione
  - ▶ l'ordinamento tra le righe è irrilevante
  - ▶ l'ordinamento tra le colonne è irrilevante

Squadra di Casa	Squadra Ospitata	Reti SC	Reti SO
Inter	Juventus	2	0
Lazio	Roma	0	0
Inter	Milan	5	0
Bari	Milan	0	1



String x String x Intero x Intero

Il modello Relazionale
- 9 -
Basi di Dati



## Relazioni e basi di dati



- In una base di dati relazionale, **le relazioni** (una o più) **sono utilizzate per organizzare i dati**
- Si possono creare **corrispondenze fra relazioni con valori comuni degli attributi**
  - ▶ In questo modo possiamo organizzare in modo estremamente efficace le informazioni provenienti dal mondo reale
- Vediamo un esempio...

Il modello Relazionale
- 10 -
Basi di Dati



## Esempio 1: Corsi ed esami



### Studenti

Matricola	Cognome	Nome	Data Nascita
276545	Rossi	Maria	25/11/1971
485745	Neri	Fabio	12/02/1972
937653	Bruni	Mario	01/12/1971

### Esami

Studente	Voto	Corso
276545	28	01
485745	25	05
937653	30	01

### Corsi

Codice	Titolo	Docente
01	Analisi	Dozio
02	Chimica	Colombo
05	Chimica	Pepe

NOTA. **Esami** crea una corrispondenza fra **Corsi** e **Studenti**

Il modello Relazionale

- 11 -

Basi di Dati



## Indipendenza fisica dei dati



- L'esempio visto mostra che il modello logico relazionale è "basato sui valori"
  - ▶ I riferimenti fra dati in relazioni diverse sono rappresentati per mezzo dei valori dei domini che compaiono nelle tuple
- In questo modo la rappresentazione logica dei dati non fa alcun riferimento alla loro rappresentazione fisica: **indipendenza fisica dei dati**
  - ▶ Ciò non valeva per i modelli reticolare e gerarchico che richiedevano puntatori

Il modello Relazionale

- 12 -

Basi di Dati



## Definizioni e notazioni (1)



---

- **Schema di relazione**
  - Un simbolo R, detto nome della relazione, e un insieme di attributi  $X=\{A1, \dots, An\}$ ; a ogni attributo è associato un dominio
  - Si indica con **R(X)**
- **Schema di base di dati**
  - Un insieme di schemi di relazione con nomi diversi
  - Si indica con  **$R=\{R1(X1), \dots, Rn(Xn)\}$**

### Esami

Studente	Voto	Corso
276545	28	01
485745	25	05
937653	30	01

Esempio

- Schema Base di Dati *Poli(Studenti, Corsi, Esami)*
- Schema Relazione Esami *Esami(Studente, Voto, Corso)*

---

Il modello Relazionale
- 13 -
Basi di Dati



## Definizioni e notazioni (2)



---

- **Istanza di relazione** su uno schema R(X) è il suo "contenuto"
  - Insieme r di tuple su X
- **Istanza di base di dati** su uno schema  $R= \{R1(X1), \dots, Rn(Xn)\}$ 
  - Insieme di relazioni  $r = \{r1, \dots, rn\}$  (con ri relazione su Ri)
- **Valori degli attributi** di una tupla
  - Se t è una tupla su X e  $A \in X$ , allora t[A] (o t.A ) indica il valore di t su A

### Esami

Studente	Voto	Corso
276545	28	01
485745	25	05
937653	30	01

Esempio

- Istanza Relazione Esami  $\{(276545,28,01), (485745,25,05), (937653,30,01)\}$
- Se  $t=(276545,28,01)$  allora
  - t[Studente]=(276545)
  - t[Voto,Corso]=(28,01)

---

Il modello Relazionale
- 14 -
Basi di Dati



## Esempio 1



R={STUDENTI(Matricola, Cognome, Nome, Data di Nascita),  
ESAMI(Studente, Voto, Corso),  
CORSI(Codice, Titolo, Docente)}

- NOTE
  - ▶ La definizione consente anche relazioni con un solo attributo

### Studenti

Matricola	Cognome	Nome	Data Nascita
602495	Colombo	Cesare	29/08/71
612835	Negri	Dario	22/07/73
672231	Rossi	Mario	12/12/80

### Lavoratori

Matricola
602495
672231

Il modello Relazionale

- 15 -

Basi di Dati



## Esempio 2: Ricevute fiscali (1)



"DA MARIO"		
Ricevuta n. 2357		
Del 5/5/92		
3	Coperti	6000
2	Antipasti	12000
3	Primi	27000
2	Bistecche	36000
Totale		81000

"DA MARIO"		
Ricevuta n. 2334		
Del 4/7/92		
2	Coperti	4000
1	Antipasti	6000
2	Primi	15000
2	Orate	50000
2	Caffé	3000
Totale		78000

- Come rappresento le ricevute in un DBMS relazionale?... Si nota che la struttura non è fissa, perciò...

Il modello Relazionale

- 16 -

Basi di Dati



## Esempio 2: Ricevute fiscali (2)



### Ricevute

Num	Data	Totale
2357	5/5/1992	81000
2334	4/7/1992	78000

### Dettaglio

Num	Qta	Descr	Importo
2357	3	Coperti	6000
2357	2	Antipasti	12000
2357	3	Primi	27000
2357	2	Bistecche	36000
2334	2	Coperti	4000
2334	1	Antipasti	6000
2334	2	Primi	15000
2334	2	Orate	50000
2334	2	caffé	3000

NOTA: va ancora bene se si vuole tenere traccia dell'ordine delle righe di ciascuna ricevuta?  
I dati da rappresentare sono diversi a seconda degli obiettivi.

Il modello Relazionale
- 17 -
Basi di Dati



## Valori nulli (1)



- Il modello relazionale è semplice e potente, ma impone un certo grado di rigidità
  - ▶ PERSONE(Cognome, Nome, Indirizzo, Telefono)
  - ▶ Come procedo se non sempre conosco il telefono?
- Come rappresento l'assenza di informazione?
  - ▶ Potrei usare un valore del dominio ma
    - non sempre esiste
    - La distinzione fra valori veri e "non veri" è nascosta, sono i programmi che devono rendersene conto
- Si estende il concetto di relazione prevedendo che una tupla possa assumere, su ciascun attributo, un valore speciale aggiuntivo detto *valore nullo*

Il modello Relazionale
- 18 -
Basi di Dati



## Valori nulli (2)



---

### Prefetture

Città	Indirizzo
Roma	Via 4 Novembre
Firenze	NULL

### Ricevute

Num	Data	Totale
2357	5/5/1992	81000
2334	4/7/1992	78000
<del>NULL</del>	4/7/1992	78000

### Dettaglio

Num	Qta	Descr	Importo
2357	3	Coperti	6000
<del>NULL</del>	2	Antipasti	12000

**NOTA: su alcuni attributi DEVO EVITARE valori nulli!**

---

Il modello Relazionale
- 19 -
Basi di Dati



## Vincoli di Integrità



---

- Proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette per l'applicazione
- Possiamo classificare i vincoli a seconda degli elementi di una base di dati che coinvolge
  - ▶ Vincoli *intrarelazionali*: il loro soddisfacimento è definito all'interno di una singola relazione
    - Esempio: il *Voto* di *ESAMI* deve essere inferiore a 30
  - ▶ Vincoli *interrelazionali*: coinvolgono più relazioni
    - Esempio: Un numero di *Matricola* può comparire nella relazione *Esami* solo se c'è nella relazione *Studenti*
- Nel seguito vediamo le tipologie principali: vincolo di tupla, chiave, vincolo integrità referenziale

---

Il modello Relazionale
- 20 -
Basi di Dati



## Vincoli di tupla



- Esprimono condizioni sui valori di ciascuna tupla indipendentemente dalle altre tuple
- Possono essere espressi con espressioni booleane: le tuple che non verificano la condizione non vengono accettate dal DBMS
- Esempio

### ESAMI

Matricola	Voto	Lode	Corso
602495	36	null	Analisi
612835	28	lode	Chimica
672231	30	lode	Analisi 2

- ▶  $(Voto \geq 18) \text{ AND } (Voto \leq 30)$
- ▶  $!(Lode = "lode" \text{ and } (Voto <> 30))$



## Chiave (1)



- Una **chiave** è un particolare insieme di attributi che identificano univocamente le tuple di una relazione (definizione intuitiva)
- Esempio
  - ▶ STUDENTI(Matricola, Codice Fiscale, Nome, Cognome, Nascita, Corso)
  - ▶ Alcune possibili chiavi possono essere  $\{Matricola\}$ ,  $\{Codice Fiscale\}$
  - ▶ Si noti che esiste sempre almeno una chiave, al più è costituita da TUTTI gli attributi
    - Una relazione è un insieme: non possono esserci due righe uguali!
  - ▶ Domanda:  $\{Matricola, Nome\}$  è chiave?



## Chiave (2)



---

- Definizione formale
  - ▶ Un insieme K di attributi è **superchiave** di una relazione r se r non contiene due tuple distinte t1 e t2 con t1[K]=t2[K]
  - ▶ **K è chiave se è una superchiave minimale di r**
    - non esiste un'altra superchiave che sia contenuta in K come sottoinsieme proprio
- Esempio: STUDENTI(Matricola, Nome, Cognome, Nascita, Corso)
  - ▶ {Matricola, Nome, Cognome}: E' superchiave, ma NON minimale (infatti {Matricola} è superchiave), quindi NON è chiave
  - ▶ {Matricola}: E' superchiave minimale (è 1 solo attributo) e quindi è chiave
- Attenzione: La chiave è un vincolo!
  - ▶ Guardo il mondo reale e trovo la chiave, quando la fisso diventa un vincolo sullo schema (quindi devono essere verificate da tutte le istanze)

---

Il modello Relazionale

- 23 -

Basi di Dati



## Chiave (3)



---

- L'esistenza di **almeno una chiave** garantisce:
  - ▶ L'accessibilità a tutti i valori di una base di dati
  - ▶ La loro univoca identificabilità e perciò la possibilità di creare corrispondenze fra relazioni
- Esempio
  - ▶ E' grazie al fatto che Codice è chiave di corso che posso creare una corrispondenza fra le tabelle Esami e Corsi

**Esami**

Studente	Voto	Corso
276545	28	01

**Corsi**

<u>Codice</u>	Titolo	Docente
01	Analisi	Dozio



---

Il modello Relazionale

- 24 -

Basi di Dati



## Chiavi e valori nulli



<u>Matricola</u>	Cod_Fis	Cognome	Nome
NULL	NULL	Rossi	Mario
602495	NULL	Neri	Luca
NULL	NRILC...	Neri	Luca

- Chiavi: {Matricola} e {Cod\_Fis}
- La prima riga ha valori nulli su tutte le chiavi, le righe non sono identificabili e non posso fare corrispondenze
- Le ultime due righe sono identificabili da una delle due chiavi, ma i due studenti non sono comunque distinguibili
- Per evitare questi problemi **SU UNA DELLE CHIAVI SI VIETANO VALORI NULLI: tale chiave è detta **CHIAVE PRIMARIA**** (si indica con la sottolineatura)

---

Il modello Relazionale
- 25 -
Basi di Dati



## Integrità referenziale (1)



- Un **vincolo di integrità referenziale** (foreign key) fra un insieme di attributi X di una relazione R1 e un'altra relazione R2 è soddisfatto se i valori su X di ciascuna tupla dell'istanza R1 compaiono come valori della chiave primaria dell'istanza R2

**Esami**

<u>Studente</u>	Voto	Corso
276545	28	01
485745	25	05
937653	30	01

**Corsi**

<u>Codice</u>	Titolo	Docente
01	Analisi	Dozio
02	Chimica	Colombo
05	Chimica	Pepe

Si deve definire un vincolo fra Corso e Codice

---

Il modello Relazionale
- 26 -
Basi di Dati



## Integrità referenziale (2)



Codice	Data	Prov	Numero
100001	10/1/2001	LC	B07347
100002	12/1/2001	CO	394769

Auto

Prov	Numero	Cognome	Nome
LC	B07347	Colombo	Cesare
CO	394769	Colombo	Max

Si deve definire un vincolo fra gli attributi Provincia e Numero della tabella Infrazioni con Provincia e Numero della tabella Auto  
 NOTA: attenzione al verso delle frecce!

Il modello Relazionale
- 27 -
Basi di Dati



## Esercizio



• Rappresentare per mezzo di una o più relazioni le informazioni contenute nell'orario delle partenze da una stazione ferroviaria: numero, orario, destinazione finale, categorie, fermate intermedie, di ciascuno dei treni in partenza

Numero	Ora	Destinazione	Categoria
--------	-----	--------------	-----------

Fermate

Treno	Stazione	Ora
-------	----------	-----

Partenze

Il modello Relazionale
- 28 -
Basi di Dati



## Operazioni di interrogazione



- Abbiamo fino a questo punto visto come sono organizzate le informazioni in una base di dati relazionale
- Per completare il modello relazionale è necessario definire le operazioni
  - ▶ Il fondamento matematico/intuitivo ci consente di definire in modo semplice anche le operazioni sui dati
  - ▶ **Operazioni di aggiornamento**: funzione che, data una istanza di base di dati, produce un'altra base di dati, sullo stesso schema
  - ▶ **Operazioni di interrogazione**: funzione che, data una base di dati, produce una relazione su un dato schema
- Fra tutte le operazioni, ne esamineremo a livello teorico solo le principali operazioni di interrogazione per poter comprendere al meglio il linguaggio SQL



## Operazione 1: SELECT



- Esprime una selezione orizzontale su una tabella, cioè genera in uscita una parte delle righe della tabella originale


Studente	Voto	Corso
276545	28	01
485745	25	05
937653	30	01



Esami con Voto=30

Studente	Voto	Corso
937653	30	01

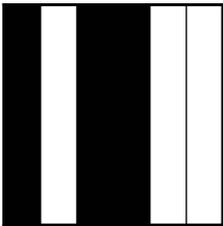


## Operazione 2: PROJECT



---

- Esprime una selezione verticale: genera in uscita una tabella con un numero ridotto di attributi



Codice	Titolo	Docente
01	Analisi	Dozio
02	Chimica	Colombo
05	Chimica	Pepe

↓

Nomi di tutti i docenti

Docente
Dozio
Colombo
Pepe

---

Il modello Relazionale
- 31 -
Basi di Dati



## Operazione 3: JOIN



---

- Esprime tutte le combinazioni fra le righe di due tabelle che verificano una particolare condizione
  - ▶ Il caso più semplice è il NATURAL JOIN: condizione di uguaglianza fra attributi eliminando i campi uguali
- In pratica
  - ▶ E' un sottoinsieme del prodotto cartesiano fra due insiemi
  - ▶ Si prendono tutte le combinazioni "sense" fra le righe di due tabelle

---

Il modello Relazionale
- 32 -
Basi di Dati



## Esempio di operazione (1)



---

- Individuare nomi e cognomi di tutti gli studenti che hanno preso 28 agli esami

**Studenti**

Matricola	Cognome	Nome
276545	Rossi	Maria
485745	Neri	Fabio
937653	Bruni	Mario

**Esami**

Studente	Voto	Corso
276545	28	01
485745	28	05
937653	30	01

---

Il modello Relazionale
- 33 -
Basi di Dati



## Esempio di operazione (2)



---

- Innanzitutto SELEZIONE dei risultati con Voto=28

**Esami**

Studente	Voto	Corso
276545	28	01
485745	28	05

- A questo punto devo trovare i nomi... Uso l'operatore di JOIN ponendo come attributi uguali **Matricola** e **Studente**...

---

Il modello Relazionale
- 34 -
Basi di Dati



### Esempio di operazione (3)

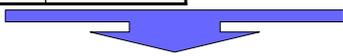


#### Studenti

Matricola	Cognome	Nome
276545	Rossi	Maria
485745	Neri	Fabio
937653	Bruni	Mario

#### Esami

Studente	Voto	Corso
276545	28	01
485745	28	05



Matricola	Cognome	Nome	Studente	Voto	Corso
276545	Rossi	Maria	276545	28	01
276545	Rossi	Maria	485745	28	05
485745	Neri	Fabio	276545	28	01
485745	Neri	Fabio	485745	28	05
937653	Bruni	Mario	276545	28	01
937653	Bruni	Mario	485745	28	05

Il modello Relazionale
- 35 -
Basi di Dati



### Esempio di operazione (4)



Matricola	Cognome	Nome	Voto	Corso
276545	Rossi	Maria	28	01
485745	Neri	Fabio	28	05

- Con una proiezione scopro che Rossi Maria e Neri Fabio sono gli studenti che hanno preso 28
- Si intuisce che complicando e combinando a piacere le interrogazioni è possibile utilizzare in modo “produttivo” la base di dati

Il modello Relazionale
- 36 -
Basi di Dati



## Commento all'esempio



- Confronta due possibili soluzioni
  - ▶ Selezione→Join→Proiezione
  - ▶ Join→Selezione→Proiezione
- Ci sono differenze?
  - ▶ Notiamo che la seconda soluzione è meno efficiente (ce lo dice la teoria matematica alla base del modello relazionale)
  - ▶ Quindi: l'ordine delle operazioni è importante!
- Per evitare problemi di questo tipo esistono linguaggi DICHIARATIVI che consentono di specificare le operazioni da fare senza specificare l'ordine (sarà il DBMS a decidere come implementarle)
- Il linguaggio più diffuso è l'**SQL**



## Cos'è SQL? (1)



- **SQL** (Structured Query Language) è il linguaggio di riferimento per le basi di dati relazionali
  - ▶ **Data Definition Language** (definizione schema)
  - ▶ **Data Manipulation Language** (interrogazione e manipolazione)
  - ▶ **Data Control Language** (controllo e sicurezza)
- Come funziona SQL?
  - ▶ SQL esprime i comandi DICHIARANDO l'obiettivo e non il modo per raggiungerlo
  - ▶ Tale comando viene passato ad un interprete che lo analizza e formula un comando equivalente utilizzando il linguaggio procedurale dello specifico DBMS



## Cos'è SQL? (2)



- Possiamo classificare le istruzioni SQL in
 

Comandi SQL (Structured Query Language)		
DDL	DML	DCL
CREATE	INSERT	GRANT
	UPDATE	REVOKE
	DELETE	
	SELECT	
- **NOTA.** Anche se sono diffuse interfacce di tipo grafico, la conoscenza di SQL rimane molto importante per lavorare con i DB (soprattutto per chi deve sviluppare SW)

Il modello Relazionale
- 39 -
Basi di Dati



## Esempio di operazione



Nome	Età	Reddito
Angelo	27	48
Marco	26	45
Piero	19	50

**Persone**

Individuare nomi e redditi delle persone che guadagnano più di 45

```

select Nome, Reddito
from Persone
where Reddito > 45;
    
```

Proiezione

Selezione

➤

Nome	Reddito
Angelo	48
Piero	50

Il modello Relazionale
- 40 -
Basi di Dati



## Bibliografia



- Paolo Atzeni, Stefano Ceri,  
Stefano Paraboschi, Riccardo Torlone  
***Basi di Dati - Seconda edizione***  
***settembre 1999 - McGraw Hill***
  - ▶ Capitolo 2

