



UNIVERSITÀ
DEGLI STUDI DI TRIESTE



Corso di Laurea in Tecniche di Radiologia Medica per immagini e Radioterapia Informatica Medica

2CFU – 20 ore

IL MODELLO DEI DATI RELAZIONALE

Prof. Sara Renata Francesca Marceglio

Modello dei dati: il modello logico

- Sono i modelli dei dati effettivamente presenti nei Sistemi di Gestione di Basi di Dati (DBMS), cioè che possono essere implementati in calcolatori (modello relazionale, modello gerarchico, modello reticolare, modello a oggetti)
- Sono chiamati “logici” perché le strutture usate da questi modelli, pur astratte, seguono una particolare organizzazione:
 - **modello relazionale → tabelle**
 - modello gerarchico → alberi
 - modello reticolare → grafi
 - modello a oggetti → oggetti

Introduzione

- Il modello relazionale dei dati si basa sui concetti di **relazione** e di **tabella**.
- Il modello relazionale dei dati rappresenta una base di dati come una collezione di **tabelle** legate tra loro.
- Ogni **tabella** corrisponde, nella trattazione teorica di tale modello di dati, al concetto matematico di **relazione**, più precisamente ad un concetto ad esso abbastanza simile:
 1. Una tabella è composta da **righe** e **colonne**;
 2. Una Relazione è un insieme di **tuple**, composte a loro volta da un insieme di valori di **attributi**.

Proprietà di una relazione

1. Non è stabilito alcun criterio di ordinamento fra le tuple di una relazione
2. Non possono esistere due tuple identiche in una relazione, in quanto non possono esistere due elementi identici in una relazione
3. Non esiste ordinamento tra gli attributi di una relazione

Relazioni, tuple, attributi

Modello Relazionale

Query Language

Relazione ↔ Tabella

Tupla ↔ Riga

Attributo ↔ Colonna

Relazione

Attributo

PATIENT	PHID	FirstName	LastName	Encounter Date	Therapy
	000ZZ000	John	Smith	2003-03-12	Flutamide
	111AA222	Mary	Brown	2004-10-14	Penicillin
	000EE999	Kevin	Green	2001-09-23	Leuprolide
	123XX456	Ann	Black	2002-05-11	Epinephrine

Tupla

Dominio

- Dominio è l'insieme dei valori ammissibili **atomici** che un attributo può assumere.
- L'atomicità consiste nel fatto che tali valori, all'interno del modello relazionale dei dati, non sono decomponibili in sottoparti (tale proprietà si esprime anche dicendo che le relazioni sono in **Prima Forma Normale**).
- Esempi:
 - Attributo **Cognome**: il dominio si esprime con **dom(Cognome)**, e consiste nell'insieme di tutti i cognomi esistenti
 - Attributo **Indirizzo di residenza**: è atomico?

Schema della Relazione

Schema della Relazione (componente **intensionale**):

Riga di intestazione della relazione, composta dal

- Nome della **Relazione** (nell'esempio **PATIENT**)
- Nomi degli **attributi** (nell'esempio **PHID, FirstName, LastName, EncounterDate, Therapy**)

PATIENT	PHID	FirstName	LastName	Encounter Date	Therapy
	000ZZ000	John	Smith	2003-03-12	Flutamide
	111AA222	Mary	Brown	2004-10-14	Penicillin
	000EE999	Kevin	Green	2001-09-23	Leuprolide
	123XX456	Ann	Black	2002-05-11	Epinephrine

Istanza della Relazione

Istanza della Relazione (componente estensionale):

Le tuple (righe) successive alla intestazione della relazione (tabella), costituiscono l'istanza della relazione.

PATIENT	PHID	FirstName	LastName	Encounter Date	Therapy
	000ZZ000	John	Smith	2003-03-12	Flutamide
	111AA222	Mary	Brown	2004-10-14	Penicillin
	000EE999	Kevin	Green	2001-09-23	Leuprolide
	123XX456	Ann	Black	2002-05-11	Epinephrine

Notazione dello Schema di una relazione

- In generale lo schema di una relazione si indica con:

$$R(A_1, A_2, \dots, A_n)$$

- R = nome della relazione
 - $\text{dom}(A_i)$ è il dominio dell'attributo i -esimo
 - n è il grado della relazione (cioè il numero degli attributi)
- Lo schema di una relazione si indica anche con $R(X)$ dove
 $X = \{A_1, A_2, \dots, A_n\}$ cioè l'insieme X degli attributi

- Esempio:

PATIENT (PHID, FirstName, LastName, EncounterDate, Therapy)

Notazione dello Schema di una Base di dati

- Uno schema di Base di dati è un insieme di schemi di relazione, si indica con:

$$\mathbf{R} = \{ R_1(X_1), R_2(X_2), \dots, R_n(X_n) \}$$

- R = nome del database
- R_i = nome delle singole relazioni

Il concetto di chiave

- Una Relazione è un insieme di tuple in cui essendo per definizione distinti tutti gli elementi di un insieme, ***non possono esistere all'interno di una stessa relazione due tuple identiche***, aventi cioè la stessa combinazione di valori per tutti gli attributi.
- Per una relazione deve quindi esistere un sottoinsieme di attributi che conservano la **proprietà** per cui **non esistono due o più tuple con la stessa combinazione di valori per questi attributi → CHIAVI**

Superchiavi

PATIENT	PHID	FirstName	LastName	BirthDate	BirthPlace	GP	Diagnosis
	000ZZ000	John	Smith	1980-03-12	New York	Parker	Diabetes
	080JJ333	John	Smith	1945-11-08	Los Angeles	Jackson	Hepatitis
	111AA222	Mary	Brown	1955-10-14	San Antonio	Hart	Hypertension
	000EE999	Kevin	Green	1974-09-23	Sydney	Goldman	Cold
	123XX456	Ann	Black	1963-05-11	Frankfurt	O'Neill	Miocarditis

- Si definisce **Superchiave** ogni sottoinsieme di attributi di una relazione che soddisfa la proprietà di **non avere la medesima combinazione di valori degli attributi per due o più tuple**
- Una **Superchiave** è un insieme di attributi che consentono in una relazione di **individuare univocamente** una tupla.
- Possono esistere **più superchiavi di una stessa relazione**

Chiave

- Si dice **chiave** una **superchiave minimale**, ovvero una superchiave dalla quale non è possibile individuare un suo sottoinsieme proprio di attributi, che conservi la proprietà di essere superchiave.
- La Superchiave {FirstName, LastName, BirthDate, BirthPlace, GP} NON è chiave della relazione: è infatti possibile estrarre il sottoinsieme di attributi {FirstName, LastName, BirthDate, BirthPlace} che conserva ancora la proprietà di essere chiave
- L'insieme di attributi {FirstName, LastName, BirthDate, BirthPlace}, a sua volta, è una superchiave minimale e, dunque, chiave.
- L'insieme di attributi {PHID}, superchiave, è evidentemente minimale ed è dunque un'altra chiave della relazione.

Chiave Primaria

- In generale esistono più chiavi per una relazione: si parla in questo caso di *candidati chiave*.
- Si definisce **Chiave Primaria** la chiave scelta tra i candidati chiave per identificare le tuple della relazione considerata. La chiave primaria non può mai essere nulla
- Convenzione adottata per identificare la chiave primaria di una relazione consiste nel far seguire il termine % al nome degli attributi che costituiscono la chiave primaria.

Chiave Primaria - Esempi

1) Adottiamo la chiave primaria

{FirstName, LastName, BirthDate, BirthPlace}

Per la relazione PATIENT

PATIENT (PHID, FirstName%, LastName%, BirthDate%, BirthPlace%, GP, Diagnosis)

2) Adottiamo la chiave primaria

{PHID}

per la relazione PATIENT

PATIENT (PHID%, FirstName, LastName, BirthDate, BirthPlace, GP, Diagnosis)

Vincoli

- In generale non è vero che qualsiasi insieme di tuple sullo schema rappresenti informazioni corrette per l'applicazione → si introduce il concetto di **vincolo di integrità**, come proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette.
- Un vincolo è **intra-relazionale** se il suo soddisfacimento è definito rispetto a singole relazioni della base di dati;
 - I **vincoli di tupla** esprimono condizioni sui valori di ciascuna tupla indipendentemente dalle altre (es. voto di esami universitari sufficienti nell'intervallo [18, 30] e 30 e lode); attributo NOT NULL.
 - I **vincoli di chiave**, soddisfano la univoca identificabilità delle tuple di una relazione.
- Un vincolo è **inter-relazionale** se coinvolge più relazioni;
 - **Vincoli di integrità referenziale**

Vincoli di integrità referenziale

- Un **Vincolo di integrità referenziale** è basato sul concetto di **chiave esterna** (foreign key)
- **Esiste un vincolo di integrità referenziale** fra un insieme di attributi X di una relazione R_1 e un'altra relazione R_2 se i valori su X di ciascuna tupla dell'istanza di R_1 compaiono come valori della chiave (primaria) dell'istanza di R_2 .

$R_1 (A_{1,1}\%, A_{1,2}, \dots, A_{1,FK}, \dots)$

$R_2 (A_{2,1}\%, A_{2,2}, \dots, A_{2,N}, \dots)$



Esempio

Chiave esterna

PRESCRIPTIONS	Patient	Operative Unit	Doctor	Drug name
	1	3	1	Paracetamol
	3	2	1	Antibiotics
	1	3	2	Melatonin

1. I due attributi devono avere lo stesso dominio
2. I valori che assume la chiave esterna possono essere solo quelli esistenti nella chiave primaria, oppure possono essere nulli

OPERATIVE UNIT	Unit_Number	Name	Specialty	n.beds
	1	Cardiology 1	Cardiology	55
	2	G.Washington	Oncology	37
	3	M. Montessori	Pediatrics	47

Chiave primaria

Il valore *null* degli attributi

- 1- Non valido per l'istanza corrente. Ad esempio cognome da coniugato per un paziente maschio.
- 2- Valido ma non ancora esistente. Ad esempio cognome da coniugata di una paziente nubile.
- 3- Esistente ma la cui memorizzazione non è consentita a livello logico. Ad esempio la religione di un paziente, che in alcuni paesi (come gli USA) non è dichiarata per evitare discriminazioni.
- 4- Esistente ma non noto. Ad esempio informazioni acquisite sul paziente durante una degenza in un altro Ospedale.
- 5- Esistente, ma non ancora memorizzato. Ad esempio la anamnesi di un paziente che non è ancora stata raccolta perché il soggetto è stato appena ricoverato.
- 6- Memorizzato ma poi cancellato. Ad esempio perché l'informazione è stata inserita erroneamente.
- 7- Disponibile ma in fase di aggiornamento. Ad esempio la terapia assegnata ad un paziente, che in seguito viene modificata.
- 8- Disponibile ma di validità sospetta. Ad esempio la diagnosi di una certa patologia in contrasto con alcuni valori di esami che documentano la assenza di tale patologia.
- 9- Disponibile ma non valido. Ad esempio il valore di un parametro ematochimico ben superiore ai massimi consentiti anche in casi patologici.
- 10- Derivato da un altro dato che a sua volta assume valore nullo. Ad esempio la superficie corporea di un paziente (BSA³) di cui non sia stata rilevata l'altezza.