

Introduzione alle Basi di Dati

Master in Sistemi Informativi Geografici per il Monitoraggio e la
Gestione del Territorio

Claudio Rocchini
claudio@rockini.name

Istituto Geografico Militare

2012

Sommario

Vediamo brevemente alcuni aspetti che riguardano i database, con particolare riferimento al loro utilizzo nei GIS.

Una introduzione (anche breve) alle Basi di Dati, richiederebbe un intero corso: si presenteranno i concetti minimi necessari alla comprensione delle basi di dati utilizzate nei sistemi geografici.

Dipendenti IGM

Obiettivo

Si vuole realizzare una base di dati per gestire il personale dell'IGM.

Realizzazione

- Per ogni dipendente devono essere memorizzati i dati personali: nome, cognome, data nascita;
- si vuole tenere un elenco degli uffici dell'IGM, con i relativi dati associati;
- ogni dipendente deve essere assegnato ad un ufficio;

Cartografia 1:25000

Obiettivo

Si vuole realizzare una base di dati per memorizzare la rete stradale italiana.

Realizzazione

- Per ogni strada saranno memorizzate le informazioni necessarie;
- i dati saranno suddivisi in fogli cartografici: ogni strada sarà contenuta in un foglio;
- sarà creata una lista di categorie stradali: ogni strada farà parte di una categoria.

Il Modello Relazionali

Le basi di dati moderne sono rappresentate dal cosiddetto **Modello Relazionale**. In passato esistevano altri modelli (gerarchico, reticolare).

Definizione del Modello Relazionale

- Entità** sono gli oggetti distinguibili del nostro database, sono identificati dagli attributi (es. dipendenti, uffici, fogli, strade);
- Relazioni** legano fra loro le entità. Possono essere di tre tipi: uno a uno, uno a molti (o molti a uno), molti a molti (es. appartiene, é della categoria di);
- Attributi** sono le proprietà delle entità (cognome, nome, nome strada).

Tipi di relazione

Uno a Uno

Capi Ufficio verso Uffici: ad ogni capo ufficio corrisponde un ufficio e viceversa.

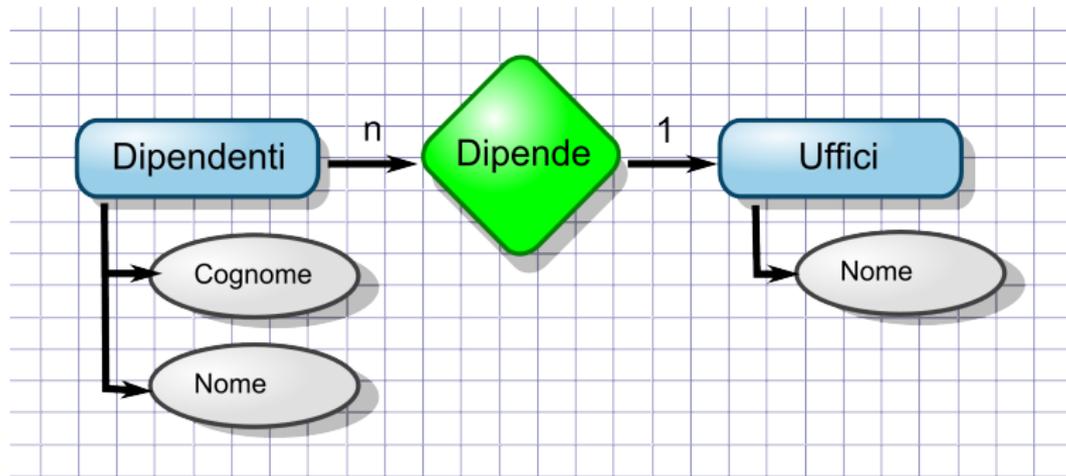
Uno a Molti

(La piú diffusa) Impiegati verso Uffici: ad ogni impiegato corrisponde un ufficio, ma ad ogni ufficio corrispondono molti impiegati.

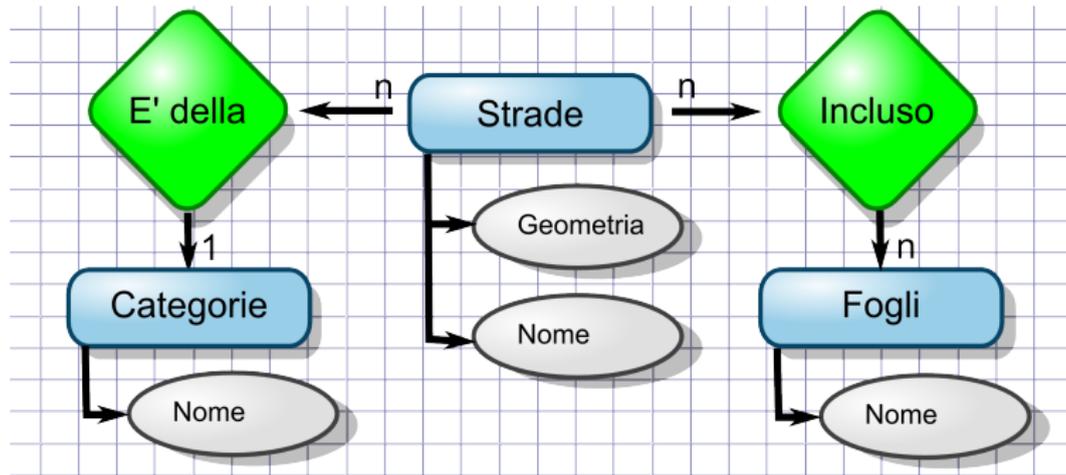
Molti a Molti

Strade verso comuni: una strada passa per molti comune, mentre per un comune passano molte strade.

Schema ER amministrativo



Schema ER geografico



Le tabelle

- Il **modello relazionale** é realizzato in pratica attraverso una serie di tabelle.
- Ogni entità é associata ad una tabella di dati;
 - le **righe** rappresentano gli elementi delle entità (ogni singolo dipendente);
 - le **colonne** rappresentano gli attributi (nome, cognome,...).
- Le relazioni sono realizzate:
 - **implicitamente** tramite la corrispondenza dei valori di opportune colonne (uno a uno, uno a molti);
 - **esplicitamente** tramite una tabella di corrispondenza apposita (molti a molti).
- Nota: non c'è un ordine prestabilito nelle righe di una tabella.

Relazione uno a molti tabellare

Strade

Id	Nome	Foglio	Categoria
1	Aurelia	645NE	B7
2	Emilia	645NE	A5
...
45	SS67	324SO	B4

Classifica

Id	Descrizione
A5	Autostrada
B4	Strada Urbana
...	...
B7	Strada Secondaria

Le relazioni sono realizzate tramite campi chiave di riferimento: la colonna classifica di Strade corrisponde alla colonna Id di Classifica.

Proprietá di un Documento

Esempi di documenti

- Le lettere scritte con Word.
- Le fotografie manipolate con Photoshop.

Proprietá

- Creazione di uno vuoto.
- Caricamento documento da disco in memoria computer.
- Operazioni di modifica in memoria.
- Salvataggio (oppure no) da memoria a disco.
- Manipolazione del file (cancellazione, copia, spedizione, ...).

Proprietá di un Database

Esempi di database

- I dati amministrativi di un Ente Pubblico.
- I dati cartografici vettoriali.

Proprietá

- Lettura dei dati tramite connessione senza caricamento).
- Registrazione istantanea delle modifiche.
- Non c'è il concetto di caricamento/salvataggio (i dati sono sempre su disco).
- Non c'è associazione con file (no copia, cancellazione, diretta).
- Utilizzo condiviso fra piú utenti contemporanei.

Tipi degli Attributi

Ogni colonna della tabella di un prefissato **tipo**.

I valori degli attributi del database sono:

Testo Con lunghezza prefissata oppure no.

Numeri Con varie opzioni, es. numero decimali od interi.

Date Oppure orari della giornata od intervalli temporali.

Valori di verità vero o falso (chiamati boolean).

Indici automatici Numeri interi che incrementano da soli.

Campi Speciali dati binari generici oppure dati geometrici.

Esempi di tipo degli attributi

- Nome e Cognome, descrizioni: testo.
- Date di Nascita o di Costruzione: data.
- Numero di corsie: numero intero.
- Stipendio: Numero con virgola e due decimali.
- Coniugato: valore di verità (vero o falso).
- Forma di un edificio, immagine: campo binario o geometrico.

Chiavi e Chiavi Primarie

Chiavi

Una **chiave** é un attributo (od un insieme di attributi) utilizzato nelle relazioni fra tabelle.

Chiave Primaria

La **Chiave Primaria** é un attributo (od un insieme di attributi) che identifica univocamente un oggetto, vale a dire una riga di una tabella.

Esempi

Nella tabella Dipendente la chiave primaria pué essere l'insieme NOME + COGNOME + DATA_NASCITA. Il codice fiscale un esempio di chiave primaria utilizzata per identificare le persone.

Indici sui dati

- Gli **indici** sono strutture aggiuntive che permettono di velocizzare le interrogazioni sui campi dati.
- Un indice é di solito associato ad un attributo (colonna di una tabella), oppure ad un insieme di colonne.
- Se si prevede di fare ricerche su di una colonna od un insieme di colonne, é opportuno costruire un indice.
- Gli indici possono rallentare la modifica o l'inserimento di dati.
- Gli indici occupano spazio disco.

Indici Spaziali

- Gli attributi di tipo geometrico possono essere associati ad **indici spaziali**.
- Gli indici spaziali sono fondamentali per il funzionamento dei database cartografici.

Esempi

Geomedia usa la colonna Geometry_SK per costruire un indice spaziale.

I database cartografici professionali (Oracle,Postgres), possono costruire i loro indici spaziali.

Schemi

Scopo

I database reali possono essere molto complessi e contenere migliaia di oggetti.

Definizione

Gli **schemi** rappresentano la possibilità di suddividere il database in sottoparti. Ogni oggetto del database apparterrá ad uno schema.

Esempi

In Oracle ad ogni utente é associato un omonimo schema.
Il database geodetico IGM é diviso negli schemi: punti trigonometrici, livellazione, IGM95.

Transazioni

Scopo

Rendere monolitiche le operazioni complesse (su piú dati contemporaneamente).

SQL

```
BEGIN TRANSTACTION;
```

```
INSERT ...
```

```
INSERT ...
```

```
UPDATE ...
```

```
COMMIT; (oppure ROLLBACK)
```

Prima Forma Normale

Definizione

Gli attributi non devono essere composti. Non devono esistere righe duplicate (nei dati indicativi).

Controesempio

Id	Descrizione	Dati
A107	Edificio	WD1;35;A01;0;BRN;11
C103	Ponte	WD1;35;A01;0;BRN;07
C103	Ponte	WD1;35;A02;0;BRN;08

Seconda Forma Normale

Definizione

Gli attributi dipendono dall'intera chiave e non da una parte di essa.

Controesempio

Cognome	Nome	Iniz.Cognome
Bonaparte	Napoleone	N
Cesare	Giulio	G
Rocchini	Claudio	R

Terza Forma Normale

Definizione

Gli attributi non chiave devono essere reciprocamente indipendenti: devono dipendere solo dalla chiave.

Controesempio

Id	FACC	LAB	LTN
307	LAP030	L301	4
308	LAP030	L303B	2
309	LAP020	L301A	2

Query

Scopo delle Query

- Estrarre un sottoinsieme di dati che rispetta particolari condizioni.
- Creare correlazioni che permettono di **vedere** in modo diverso i dati.
- Utilizzare le relazioni presenti fra tabelle.
- I risultati delle interrogazioni possono essere utilizzati per creare nuovi dati.

Tipi di Query

Le Query possono essere:

- Sottoinsieme di attributi: si seleziona un'insieme di colonne.
- Sottoinsieme di oggetti con criteri sugli attributi: filtro sugli attributi.
- **Join** di relazioni: creazione di dati da relazioni fra tabelle.
- Altro (es. query con filtri spaziali).

Viste

Il concetto di vista

- Le query sono **viste** sui dati.
- Modificando i dati si modificano anche il risultato delle query (comportamento dinamico).
- I database professionali permettono di creare delle viste a partire da una query.
- Alcuni GIS permettono di lavorare con query sui dati.

Introduzione a SQL

Definizione

SQL é un linguaggio standard per la manipolazione dei dati. Con SQL é possibile definire ogni tipo di query. Il simbolo * vuol dire tutto. Spesso la sintassi é autoesplicitiva (basta leggerla).

Seleziona un'intera Tabella

```
SELECT *  
FROM dipendenti;
```

Esempi di filtri SQL

Selezione di alcune colonne

```
SELECT cognome, nome  
FROM dipendenti;
```

Selezione di alcune righe

```
SELECT *  
FROM dipendenti  
WHERE nome = 'claudio'  
AND anno_nascita < 1970;
```

I dati **stringa** vanno scritti fra apicetti

Esempi di Join

Join amministrativa

```
SELECT cognome.dipendenti, uffici.nome  
FROM dipendenti,uffici  
WHERE dipdendenti.ufficio = uffici.nome;
```

Join spaziale

```
SELECT strade.*, case.*  
FROM strade,case  
WHERE touch(strade.geometry,case.geometry);
```

In presenza di piú tabelle, bisogna specificare da quale tabella vengono le colonne.

Query Spaziali Geomeia

Geomeia di Intergraph

Puó eseguire alcune query spaziali: *Tocca*, *Entro la distanza di ...*, *Contiene*, *É Contenuto*, *Contiene Strettamente*, *É contenuto strettamente*, *Si sovrappone*, *Incontra*, *Equivale*. Dalla versione 5 in poi é possibile aggiungere il **non**.



Query Spaziali ArcGIS

ArcGIS di ESRI

Puó impostare alcune relazioni spaziali (regole topologiche), ecco alcuni esempi

The screenshot displays four panels, each representing a different topological rule in ArcGIS. Each panel includes a 'Rule Description' with illustrative diagrams and a dropdown menu for the rule name.

- Must Not Have Gaps:** The rule description states that a void cannot exist between areas in the same layer and that the boundary of any void that does exist is an error. The dropdown menu is set to 'Must Not Have Gaps'.
- Must Be Covered By Feature Class Of:** The rule description states that area features of one layer must cover the area features of another layer, and that any area where features in the second layer are not covered by features in the first layer is an error. The dropdown menu is set to 'Must Be Covered By Feature Class Of'.
- Contains Point:** The rule description states that area features from one layer must contain at least one point feature from another layer, and that any area feature that does not contain at least one point feature is an error. The dropdown menu is set to 'Contains Point'.
- Boundary Must Be Covered By:** The rule description states that the boundaries of area features from one layer must be covered by line features of another layer, and that the boundary of an area feature that is not coincident with a line feature is an error. The dropdown menu is set to 'Boundary Must Be Covered By'.

Query Spaziali Oracle

Oracle

Oracle Spatial ha un supporto nativo per la Cartografia.

Esempi di SQL

```
SELECT strade.*, case.*  
FROM strade.case  
WHERE SDO_RELATE(A.Geom,B.Geom,'mask=FILTER')='TRUE' ;
```

FILTER

É una combinazione dei seguenti controlli:

*OVERLAPBDYDISJOINT, OVERLAPBDYINTERSECT, EQUAL,
INSIDE, COVEREDBY, CONTAINS, COVERS, ANYINTERACT, ON.*

Indice

- 1 Esempi
 - Esempio Amministrativo
 - Esempio geografico
- 2 Il Modello Relazionale
 - Introduzione
 - Entità e relazioni
 - Realizzazione Tabellare
- 3 Documenti e Basi di Dati
 - Il concetto di documento
 - Il concetto di database
- 4 Concetti
 - I Tipo di Dato
 - Chiavi
 - Indici
 - Schemi
 - Transazioni
- 5 Forme Normali
 - Prima Forma Normale
 - Seconda Forma Normale
 - Terza Forma Normale
- 6 Interrogazioni
 - Query
 - Viste
 - SQL
 - Query Spaziali