

Corso
di
Basi di Dati Spaziali

**Modello logico
per dati spaziali**

Angelo Montanari
Donatella Gubiani

Modelli logici

- Permettono una rappresentazione astratta dei dati
- La maggior parte dei sistemi di basi di dati oggi sul mercato si basa sul modello relazionale

2

Il modello relazionale

- Proposto da E. F. Codd nel 1970 per favorire l'indipendenza dei dati
- Disponibile in DBMS reali nel 1981 (non è facile coniugare l'indipendenza con l'efficienza e l'affidabilità!)
- Si basa sul concetto matematico di relazione (con una variante)
- Le relazioni hanno una naturale rappresentazione per mezzo di tabelle

3

Definizioni e notazioni - 1

- Schema di relazione:
un nome R con un insieme di attributi $X = A_1, \dots, A_n$:
$$R(X) = R(A_1, \dots, A_n)$$
- Schema di base di dati:
insieme di schemi di relazione:
$$R = \{R_1(X_1), \dots, R_k(X_k)\}$$

4

Definizioni e notazioni - 2

- Una ennupla su un insieme di attributi X è una funzione che associa a ciascun attributo A in X un valore del dominio di A
 - $t[X]$ denota il valore della ennupla t sull'insieme di attributi X
 - $t[A]$ denota il valore della ennupla t sull'attributo A

5

Definizioni e notazioni - 3

- Istanza di relazione su uno schema $R(X)$ è un insieme r di ennuple su X
 $r(R)$ o $r \in R(X)$
- Istanza di base di dati su uno schema $R = \{R_1(X_1), \dots, R_n(X_n)\}$ è un insieme di istanze di relazioni
 $r = \{r_1, \dots, r_n\}$ (con r_i relazione su R_i)

6

Modello GEO-relazionale

- Il modello relazionale prevede una serie di tipi di dato di base:
 - INTEGER: numeri interi
 - REAL: numeri reali approssimati
 - STRING: stringhe di caratteri
 - BOOLEAN: valori booleani (vero, falso)
- Per trattare dati spaziali vengono introdotti uno o più tipi di dato spaziale (SDT)

7

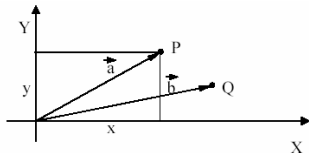
Tipi di dato spaziale (SDT)

- Geometria lineare
 - i componenti di base sono punti, segmenti di retta, spezzate, spezzate chiuse
- Si definisce uno SDT specifico per ogni dimensione
 - dimensione 0 → punto
 - dimensione 1 → spezzata
 - dimensione 2 → poligono

8

Elementi di geometria - 1

- Spazio di riferimento: R^2
- Punto: $P(x,y)$



- \vec{a} è il vettore che individua il punto P in R^2

9

Elementi di geometria - 2

- Segmento di retta per i punti P e Q individuati dai vettori \vec{a} e \vec{b} :

$$s = \{\lambda\vec{a} + (1-\lambda)\vec{b} \mid \lambda \in [0,1]\}$$
 - chiamiamo *estremi del segmento* i punti $s.E_1=Q(\lambda=0)$ e $s.E_2=P(\lambda=1)$
- Spezzata è un insieme di segmenti di retta $sp = \{s_1, \dots, s_n\}$ tale che ogni estremo di segmento è condiviso da esattamente due segmenti, eccetto al massimo due estremi, detti *estremi della spezzata*

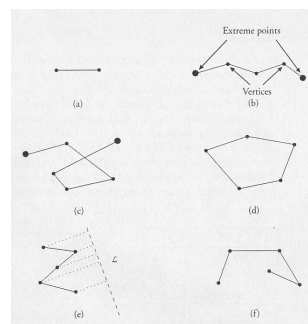
10

Elementi di geometria - 3

- se la spezzata non si interseca si dice *spezzata semplice*
- se la spezzata non ha estremi si dice *spezzata chiusa*
- una spezzata si dice *monotona* rispetto a una retta L se ogni retta L' ortogonale a L incontra la spezzata in al più un punto

11

Elementi di geometria - 4

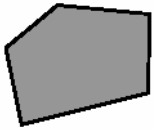


- a) segmento
- b) spezzata semplice
- c) spezzata non semplice
- d) spezzata chiusa
- e) spezzata monotona
- f) spezzata non monotona

12

Elementi di geometria - 5

- Poligono semplice: data una spezzata chiusa sp , è la porzione di R^2 delimitata da sp
 - sp fa parte del poligono



13

Elementi di geometria - 6

- Poligono bucato: data una spezzata chiusa sp e un insieme di spezzate chiuse $H = \{sp_1, \dots, sp_n\}$, è la porzione di R^2 delimitata da sp alla quale siano tolte le porzioni di piano delimitate da $sp_i \in H$, a condizione che:
 - sp e tutte le $sp_i \in H$ siano parte del poligono
 - tutte le $sp_i \in H$ siano contenute nel poligono semplice delimitato da sp
 - tutte le $sp_i \in H$ non intersechino sp
 - per ogni $sp_i, sp_j \in H$, sp_i non intersechi sp_j ed sp_i non contenga sp_j

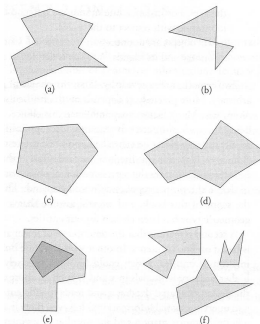
14

Elementi di geometria - 7

- Un poligono P si dice *convesso* se per ogni coppia di punti A e B in P il segmento AB è completamente incluso in P
- Un poligono monotono è un poligono *semplice* tale che la sua frontiera si può spezzare esattamente in due spezzate monotone
- Una *regione* è un insieme di poligoni

15

Elementi di geometria - 8



Esempi di poligoni:

- a) semplice
- b) non semplice
- c) convesso
- d) monotono
- e) bucato
- f) regione

16

Punto - 1

- POINT è un insieme finito P di punti di R^2
- *Boundary* di P è l'insieme vuoto
- *Interior* di P è P stesso

17

Punto - 2

- Dato P : POINT sono definite le seguenti operazioni:
 - $P.X$: REAL coordinata x del baricentro di P ;
 - $P.Y$: REAL coordinata y del baricentro di P ;
 - $P.Distanza(Q: POINT)$: REAL distanza euclidea tra il baricentro di P e il baricentro di Q ;
 - $P.NumPunti$: INTEGER numero di punti presenti in P ;
 - $P.Punto(I: INTEGER)$: POINT i -esimo punto presente in P ;
 - $P.Unione(Q: POINT)$: POINT genera l'unione dei due insiemi di punti.
- Si suppone definito anche l'operatore di uguaglianza: $P = Q$

18

Linea - 1

- LINE è un insieme di spezzate L di \mathbb{R}^2 tale che nessuna coppia di spezzate abbia come intersezione una spezzata
- *Boundary* di L è l'insieme di punti di \mathbb{R}^2 costituito dagli estremi delle spezzate tolti gli estremi che appartengono a più di una spezzata
- *Interior* di L è l'insieme di punti di \mathbb{R}^2 costituito dalle spezzate private degli estremi che costituiscono il Boundary di L

19

Linea - 2

- Data L : LINE sono definite le seguenti operazioni:
 - L .Boundary: POINT boundary di L ;
 - L .NumSpezzate: INTEGER numero di spezzate di L
 - L .Lunghezza: REAL lunghezza di L ottenuta sommando la lunghezza di tutte le spezzate che la compongono;
 - L .Semplice: BOOLEAN vero se l'unione delle spezzate di L è una spezzata semplice;
 - L .Chiusa: BOOLEAN vero se l'unione delle spezzate di L è una spezzata chiusa;
 - L .Unione(K : LINE): LINE genera un valore di tipo LINE che rappresenta l'unione di L e K ;

20

Linea - 3

- L .IntersezioneL(K : LINE): LINE genera un valore di tipo LINE che rappresenta l'intersezione di L e K . Se l'intersezione è costituita solo da punti, allora produce l'insieme vuoto;
- L .IntersezioneP(K : LINE): POINT genera un valore di tipo POINT che rappresenta l'intersezione di L e K . Se l'intersezione è costituita solo da spezzate, allora produce l'insieme vuoto. Se nell'intersezione sono presenti spezzate e punti isolati, restituisce solo questi ultimi.
- Si suppone definito anche l'operatore di uguaglianza: $L = L'$

21

Poligono - 1

- POLYGON è un insieme G di poligoni bucati di \mathbb{R}^2 , tale che ogni coppia di poligoni costituisce un insieme di punti non connesso

Definizione. Un insieme di punti si dice (path-)connesso se per ogni coppia di suoi punti esiste un cammino (path) completamente contenuto nell'insieme che li connette

- *Boundary* di G è costituito dalle spezzate che delimitano i poligoni di G
- *Interior* di G è costituito dai poligoni privati delle spezzate che li delimitano

22

Poligono - 2

- Dato G : POLYGON sono definite le seguenti operazioni (per semplicità consideriamo insiemi singoli):
 - G .Frontiera: LINE valore geometrico di tipo LINE che costituisce la frontiera esterna del poligono;
 - G .Boundary: LINE valore geometrico di tipo LINE che costituisce la frontiera del poligono;
 - G .Semplice: BOOLEAN vero se G è semplice;
 - G .NumBuchi: INTEGER numero di buchi presenti nel poligono G ;
 - G .Perimetro: REAL lunghezza del perimetro di G ottenuto sommando la lunghezza di tutte le spezzate che lo delimitano;
 - G .Area: REAL misura dell'area del poligono G ;

23

Poligono - 3

- G .Unione(H : POLYGON): POLYGON genera un valore di tipo POLYGON che rappresenta l'unione dei due poligoni G e H ;
- G .IntersezioneG(H : POLYGON): POLYGON genera un valore di tipo POLYGON che rappresenta l'intersezione dei due poligoni G e H . Se l'intersezione è costituita solo da punti e/o linee allora produce l'insieme vuoto;
- G .IntersezioneL(H : POLYGON): LINE genera un valore di tipo LINE che rappresenta l'intersezione dei due poligoni G e H . Se l'intersezione è costituita solo da punti e/o da poligoni allora produce l'insieme vuoto. Se l'intersezione presenta oltre a spezzate anche poligoni e/o punti restituisce solo le spezzate;

24

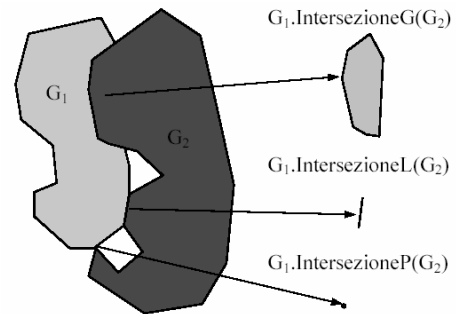
Poligono - 4

- G.IntersezioneP(H: POLYGON): POINT genera un valore di tipo POINT che rappresenta l'intersezione dei due poligoni G e H. Se l'intersezione è costituita solo da linee e/o poligoni allora produce l'insieme vuoto. Se l'intersezione presenta oltre a punti isolati anche linee e/o poligoni restituisce solo i punti isolati.

- Si suppone definito anche l'operatore di uguaglianza: $G = G'$

25

Poligono - 5



26

Relazioni (o tabelle) - 1

- Uno schema di tabella è definito come $R(X_1:D_1, \dots, X_n:D_n)$ dove D_1, \dots, D_n rappresentano domini di base o geometrici
- Un'istanza di tale schema è r di $R(X_1: D_1, \dots, X_n: D_n)$ dove $r \subseteq D_1 \times \dots \times D_n$

27

Relazioni (o tabelle) - 2

COMUNE(Nome: STRING,
NumAbitanti: INTEGER,
Estensione: POLYGON)

28

Rappresentazione dei dati spaziali: problemi

- La rappresentazione dei valori geometrici in un GEO-DBMS presenta alcuni problemi:
 - l'occupazione di memoria di un valore geometrico
 - l'imprecisione nelle coordinate (rappresentazione approssimata dei numeri reali)
 - i valori geometrici sono fortemente correlati

29

Rappresentazione dei dati spaziali: metodi

- Due metodi:
 - insieme di liste di coordinate
 - struttura basata sulla topologia dei dati

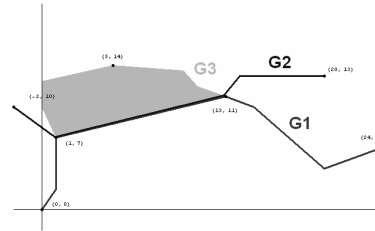
30

Liste di coordinate - 1

- Ogni tipo di dato spaziale è definito in modo analogo
 - poligono come insieme di coordinate
 - linee come insieme di coordinate
 - punti come un singoletto di coordinate

31

Liste di coordinate - 2



G1=(ln,(<0,0>,<1,2>,<1,7>,<13,11>,<15,10>,<20,4>,<24,6>))
 G2=(ln,(<-2,10>,<1,7>,<13,11>,<14,13>,<20,13>))
 G3=(pg,(<5,14>,<10,13.5>,<11,12>,<13,11>,<1,7>,<0,10>,<0,12.5>))

32

Liste di coordinate - 3

- Vantaggi:
 - l'indipendenza tra i valori geometrici
 - le buone prestazioni dell'operazione di visualizzazione grafica della geometria di un sottoinsieme limitato di valori
- Svantaggi:
 - tutte le selezioni spaziali e le operazioni di manipolazione richiedono l'applicazione di algoritmi di geometria computazionale
 - la ridondanza nella rappresentazione della geometria

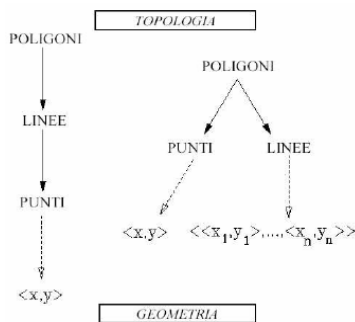
33

Struttura basata sulla topologia - 1

- Ogni tipo di dato spaziale è definito in modo diverso a seconda della sua struttura topologica
 - poligono come insieme di linee
 - linee come insieme di punti
 - punti come coppia di coordinate

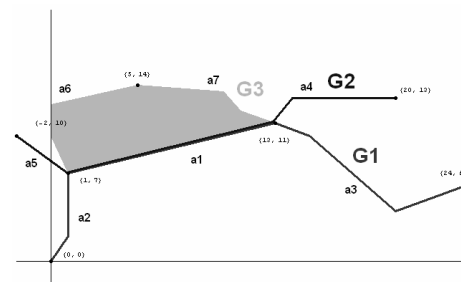
34

Struttura basata sulla topologia - 2



35

Struttura basata sulla topologia - 3



36

Struttura basata sulla topologia - 4

G1 = <ln, {a1,a2,a3}>
G2 = <ln, {a1,a4,a5}>
G3 = <pg, {a1,a6,a7}>
a1 = (P2, (), P3)
a2 = (P1, (<1,2>), P2)
a3 = (P3, (<15,10>, <20,4>), P4)
a4 = (P3, (<14,13>), P5)
a5 = (P6, (), P2)
a6 = (P2, (<0,10>, <0,12.5>), P7)
a7 = (P7, (<10,13.5>, <11,12>), P3) ³⁷

Struttura basata sulla topologia - 5

P1 = <0,0>
P2 = <1,7>
P3 = <13,11>
P4 = <24,6>
P5 = <20,13>
P6 = <-2,10>
P7 = <5,14>

38

Struttura basata sulla topologia - 6

- Vantaggi:
 - l'eliminazione della ridondanza
 - la riduzione di alcune operazioni geometriche a selezione di un insieme di identificatori o puntatori a valori geometrici elementari precalcolati
 - la riduzione della ricerca via predicati topologici ad un confronto tra identificatori o puntatori a valori geometrici elementari precalcolati
 - la verifica dei vincoli spaziali immediata e supportata dalla struttura topologica

39

Struttura basata sulla topologia - 7

- Svantaggi:
 - il ripristino della topologia corretta a valle di ogni inserimento di un nuovo valore geometrico
 - la pura visualizzazione di un valore geometrico richiede il ritrovamento della geometria attraverso una catena di puntatori (indirettezze)

40