

ANNO ACCADEMICO 2003-2004
SISTEMI INFORMATIVI GEOGRAFICI

SISTEMI INFORMATIVI TERRITORIALI (SIT)
GEOGRAPHICAL INFORMATION SYSTEMS (GIS)

Sistemi Informativi Territoriali

4. La progettazione logico/fisica di una base di dati geografica: le regole di corrispondenza tra lo schema concettuale e lo schema logico/fisico

ALBERTO BELUSSI

NOVEMBRE 2003

La progettazione logico/fisica di una Base di Dati geografica

La progettazione logico/fisica di una base di dati territoriale definisce le strutture di memorizzazione della componente spaziale e alfanumerica dell'informazione contenuta nella base stessa.

Nella progetto di una base di dati territoriale la progettazione logico/fisica costituisce una fase intermedia e segue la progettazione concettuale o fase di analisi.

Essa richiede quindi che lo schema concettuale dei dati sia già stato redatto. Lo schema dei dati è infatti il punto di partenza della progettazione logico/fisica.

L'altro essenziale prerequisito è costituito dalla scelta di un sistema per gestione di basi di dati geografiche (GEO-DBMS). In particolare, deve essere noto il modello dei dati del sistema.

Infine, poiché la strutturazione dei dati sul GEO-DBMS influisce anche sulle prestazioni dell'accesso ai dati, nella progettazione logico/fisica va tenuto presente quale sarà il tipo e la frequenza delle operazioni da eseguirsi sul contenuto della base di dati.

Come nella progettazione logica delle basi di dati tradizionali, anche per le basi di dati territoriali la progettazione logico/fisica si basa su una serie di regole o criteri generali da seguire per passare da uno schema concettuale al corrispondente schema logico.

Purtroppo nel settore delle basi di dati territoriali la situazione per quanto riguarda i modelli di riferimento non è così assestata come nelle basi di dati tradizionali. Infatti, non esiste per i GEO-DBMS l'equivalente del modello relazionale, né il modello GEO-ER risulta essere a tutt'oggi un modello standard per la progettazione concettuale.

In questo caso possiamo solo fare riferimento alla situazione che si ritiene più diffusa o più probabile, cioè, supponiamo che:

- nella progettazione concettuale si sia utilizzato il modello GEO-ER;
- il GEO-DBMS sia geo-relazionale e il suo modello dei dati sia una “specializzazione” del modello logico/fisico di riferimento presentato nella sezione 3.

Anche focalizzando su un solo modello concettuale e un solo modello logico, la progettazione logico/fisica non può seguire regole prefissate ma deve tener conto dell'uso e delle caratteristiche dei dati. Vengono quindi riportati di seguito alcuni criteri generali per la progettazione logico/fisica. Si analizzeranno poi i diversi costrutti del modello GEO-ER riportando per ciascuno diverse alternative traduzioni verso il modello logico/fisico.

I Criteri generali della progettazione logico/fisica

1. Creare una struttura normalizzata per ridurre la ridondanza dei dati e quindi le anomalie che si possono verificare in fase di aggiornamento dei dati.
2. Gestire i casi di condivisione e sovrapposizione degli elementi geometrici che rappresentano aspetti concettuali differenti.
3. Portare il più possibile le informazioni alfanumeriche in un DBMS relazionale affidabile.
4. Rappresentare il più possibile nella struttura fisica i vincoli spaziali significativi espressi nello schema concettuale.
5. Creare una struttura che sia fruibile per l'uso applicativo; questo criterio risulta ovviamente in contrasto con l'esigenza di una struttura normalizzata introdotta dal primo criterio.

Condivisione e Sovrapposizione

Condivisione: indica che due o più istanze di entità dello stesso tipo condividono materialmente lo stesso elemento geometrico nella realtà. Ad esempio, accade non di rado che due strade statali condividano un certo tratto seppur breve sul territorio.

Sovrapposizione: indica che due o più istanze di entità dello stesso tipo o di tipi diversi risultano completamente o parzialmente sovrapposte nella rappresentazione a due dimensioni pur essendo materialmente distinte sul territorio.

Ad esempio, ciò accade se due strade passano una sopra l'altra per un breve tratto, oppure nel caso di una linea che rappresenta un fiume e contemporaneamente un confine comunale.

Dallo Schema Concettuale dei Dati in GEO-ER allo schema logico/fisico su un sistema geo-relazionale

Per quanto riguarda la trasformazione dei tipi di entità normali senza attributi geometrici e dei tipi di relazioni tra entità normali valgono le regole già presentate per le basi di dati tradizionali.

Poiché il modello logico/fisico è geo-relazionale, è sempre possibile rappresentare relazioni (tabelle) in tale modello e quindi sarà possibile rappresentare in esso anche le relazioni (tabelle) che costituiscono la traduzione di tipi di entità normali e di tipi di relazioni di uno schema dei dati scritto in GEO-ER.

Per quanto riguarda invece i costrutti del modello GEO-ER che contengono riferimenti alla geometria, vengono di seguito discusse alcune proposte di traduzione nel modello logico/fisico.

D: *Come si rappresenta la geometria nel modello dei dati di un sistema geo-relazionale?*

R: *Attraverso l'introduzione nello schema logico di strati e di attributi geometrici nelle relazioni.*

D: *Quanti strati devono essere creati in una base di dati territoriale?*

D: *Quali attributi geometrici vanno messi sullo stesso strato?*

Non esiste un'unica risposta a queste domande!

Osservazioni sull'uso degli strati in uno schema logico/fisco

I dati geometrici contenuti nello stesso strato sono memorizzati in un'unica struttura dati topologica/con indice spaziale nel GEO-DBMS, ciò implica:

- la possibilità di implementare i vincoli geometrici nella struttura di dati dello strato;
- una maggiore facilità di aggiornamento dei dati contenuti nello strato, in particolare dove esistono vincoli da rispettare;
- la gestione esplicita delle condivisioni e sovrapposizioni fra i valori geometrici contenuti nello strato;
- la necessità di considerare l'intero strato per ogni accesso ai dati in esso contenuti.
- la possibilità di interrogare con maggiore efficienza i dati contenuti nello strato, in quanto le operazioni più potenti dei sistemi agiscono di solito intra-strato.

Creare strati con un gran quantità di dati geometrici richiede di solito una struttura di relazioni (tabelle) più normalizzata.

Suddivisione in strati dei dati geometrici di uno schema dei dati di GEO-ER

I dati geometrici in uno schema di GEO-ER sono descritti dai tipi di entità geometriche e dagli attributi geometrici dei tipi di entità normali.

Va quindi deciso in quanti e quali strati memorizzare i valori geometrici che rappresentano le istanze dei tipi di entità geometrica e degli attributi geometrici presenti nello schema.

Tenendo presente le osservazioni riportate in precedenza si possono considerare quindi i tre seguenti criteri:

1. porre sullo stesso strato tutte le istanze di uno **stesso tipo di entità geometrica** o dello **stesso attributo geometrico**;
2. porre preferibilmente sullo stesso strato le istanze dei tipi di entità geometriche che sono **legate da vincoli geometrici** nello schema concettuale (ad esempio, gerarchie di composizione, contenimento o intersezione non nulla);
3. porre preferibilmente sullo stesso strato le istanze dei di tipi di entità geometrica che vengono **utilizzate insieme nelle elaborazioni realizzate con maggior frequenza** dalle applicazioni.

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

TIPO DI ENTITA' NORMALE EN_1 con:

attributi alfanumerici A_1, \dots, A_n con A_1 identificatore e un attributo geometrico G_1 .

Attributi alfanumerici

Per il tipo di entità EN_1 viene creata una relazione (tabella) EN_1 che contiene tanti attributi quanti sono gli attributi alfanumerici del tipo di entità. La relazione (tabella) EN_1 ha come chiave primaria A_1 .

Attributo geometrico G_1

IPOTESI si considera il caso in cui i valori dei tipi geometrici poligono e linea non degenerino mai rispettivamente a linee e/o punti e a punti.

Si aggiunge un attributo geometrico G_1 alla relazione EN_1 con dominio geometrico scelto secondo la seguente corrispondenza:

<i>Domini GEO-ER</i>	<i>Domini del modello logico</i>
PUNTO	POINT
LINEA	LINE (+ POINT)
POLIGONO	POLYGON (+LINE, +POINT)

Se si sceglie di generare un nuovo strato per ogni tipo di entità, si genera uno strato S_{EN_1} contenente valori geometrici dello stesso dominio geometrico scelto per l'attributo G_1 .

Se invece la scelta riguardo la generazione degli strati è diversa, è possibile anche rappresentare i valori geometrici di G_1 in uno strato S preesistente.

Schema logico/fisico

Relazioni:

$EN_1(\underline{A}_1, \dots, A_n, G_1: \text{POINT/LINE/POLYGON})$

Strati

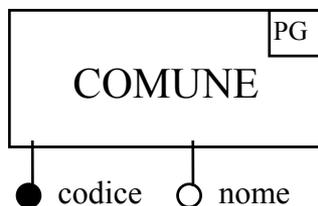
$S_EN_1(\{\text{POINT/LINE/POLYGON}\})$ oppure
 $S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$EN_1.G_1 \rightarrow S_EN_1$ oppure $EN_1.G_1 \rightarrow S$

Esempio

Schema Concettuale



Schema Logico/Fisico

Relazioni

$\text{Comune}(\underline{\text{codice}}, \text{nome}, G_1: \text{POLYGON})$

Strati

$S_Comuni: (\{\text{POLYGON}\})$ oppure
 $S(\{\text{POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$\text{Comune}.G_1 \rightarrow S_Comuni$ oppure

$\text{Comune}.G_1 \rightarrow S$

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

TIPO DI ENTITA' NORMALE EN₁ con:

**attributi alfanumerici A₁, ..., A_n, con A₁ identificatore,
un attributo geometrico G₁ di tipo LINEA (POLIGONO) e
attributi a tratti (sottoaree) T₁, ..., T_m.**

SOLUZIONE A

Schema logico/fisico

Relazioni:

EN₁(A₁, ..., A_n, G₁: LINE/POLYGON)

EN₁_TRATTI(T₁, ..., T_m, A₁, G₂: LINE/POLYGON)

Strati

S_EN₁({LINE/POLYGON}) o S({LINE/POLYGON, ...})

Mapping Attributi geometrici/Strati

EN₁.G₁ → Aggregate(S_EN₁) o Aggregate(S)

EN₁_TRATTI.G₂ → S_EN₁ o S

SOLUZIONE B

Schema logico/fisico

Relazioni:

EN₁(A₁, ..., A_n)

EN₁_TRATTI(T₁, ..., T_m, A₁, G₂: LINE/POLYGON)

Strati

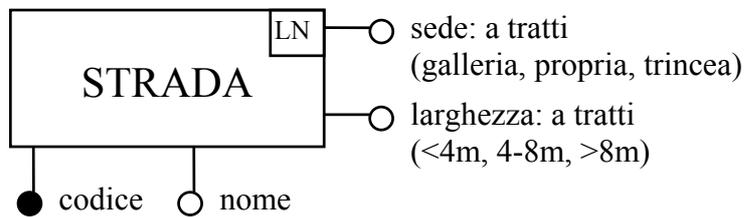
S_EN₁({LINE/POLYGON}) o S({LINE/POLYGON, ...})

Mapping Attributi geometrici/Strati

EN₁_TRATTI.G₂ → S_EN₁ o S

Esempio

Schema Concettuale



Schema logico/fisico

SOLUZIONE A

Relazioni

Strada(codice, nome, percorso: LINE)

Strada_TRATTI (sede, larghezza, codice, tracciato: LINE)

Strati

S_Strade({LINE})

Mapping Attributi geometrici/Strati

Strada.percorso → Aggregate(S_Strade)

Strada_TRATTI.tracciato → S_Strade

Istanza di STRADA: codice = SS11

GEO-ID	2	3	4	3	4	3	5	6	7
sede	G	P	P	P	P	P	T	T	G
largh.	<4	4-8	>8	4-8	>8	4-8	4-8	>8	>8

Il contenuto della base di dati nella soluzione A:

Strada

percorso	codice	nome
100	SS11	padana superiore

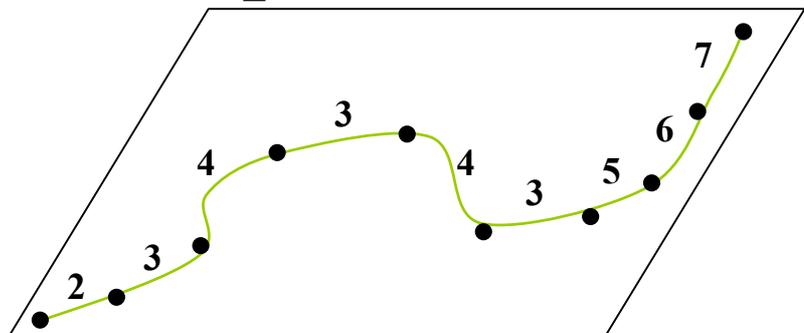
Strada TRATTI

tracciato	sede	largh	codice
2	G	<4	SS11
3	P	4-8	SS11
4	P	>8	SS11
5	T	4-8	SS11
6	T	>8	SS11
7	G	>8	SS11

Tabella aggregazione

GIDaggr	GID
100	2
100	3
100	4
100	5
100	6
100	7

S_Strade



SOLUZIONE B

Relazioni

Strada(codice, nome)

Strada_TRATTI (sede, larghezza, codice, tracciato: LINE)

Strati

S_Strade({LINE})

Mapping Attributi geometrici/Strati

Strada_TRATTI.tracciato → S_Strade

Il contenuto della base di dati nella soluzione B:

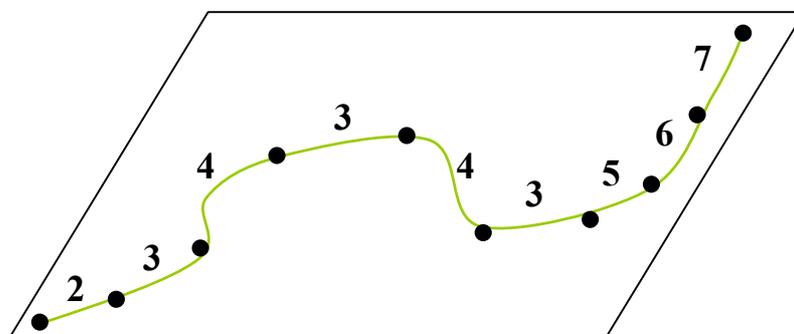
Strada

codice	nome
SS11	padana superiore

Strada TRATTI

tracciato	sede	largh	codice
2	G	<4	SS11
3	P	4-8	SS11
4	P	>8	SS11
5	T	4-8	SS11
6	T	>8	SS11
7	G	>8	SS11

S_Strade



Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

TIPO DI ENTITA' GEOMETRICA EG_1 senza attributi.

Schema logico/fisico

Relazioni:

$EG_1(G: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure
 $S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow S_EG_1$ oppure S

Attenzione: in questo caso la relazione (tabella) EG_1 contiene sempre al massimo una sola istanza.

TIPO DI ENTITA' GEOMETRICA EG_1 con attributi alfanumerici A_1, \dots, A_n .

Schema logico/fisico

Relazioni:

$EG_1(\underline{A_1}, \dots, \underline{A_n}, G: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure
 $S(\{\text{POINT/LINE/POLYGON}, \dots\})$

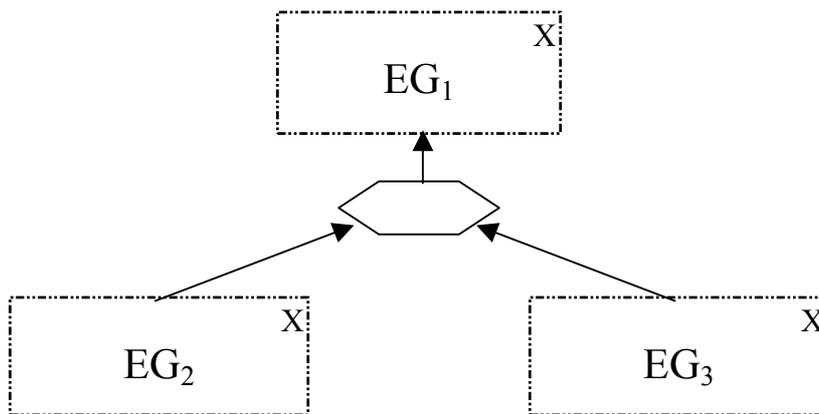
Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow S_EG_1$ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI GENERALIZZAZIONE (ISA) tra tipi di entità geometriche.

Caso A: senza attributi.



Schema logico/fisico

Relazioni:

$EG_1(\underline{\text{TIPO}}: \{EG_2, EG_3\}, G: \text{POINT/LINE/POLYGON})$

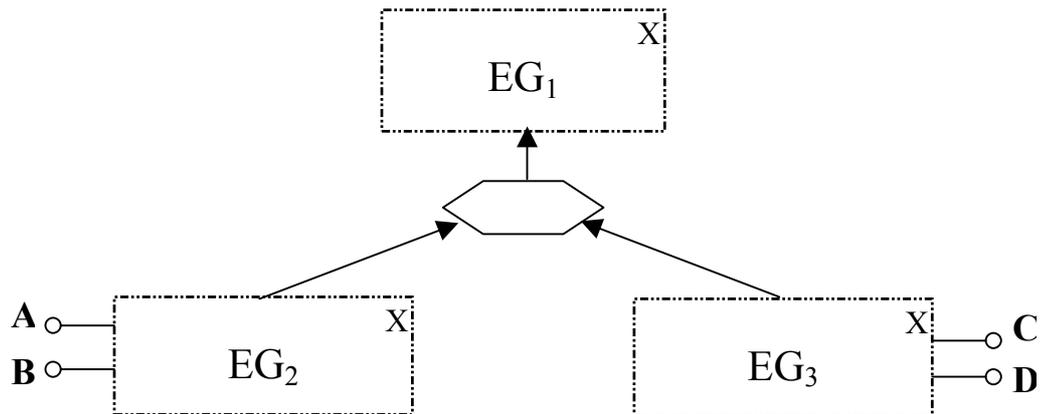
Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure
 $S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow S_EG_1$ oppure S

Caso B: con attributi.



Schema logico/fisico

Relazioni:

EG₂(A, B, G₂: POINT/LINE/POLYGON)

EG₃(C, D, G₃: POINT/LINE/POLYGON)

Strati

S_EG₁({POINT/LINE/POLYGON}) oppure

S({POINT/LINE/POLYGON,...})

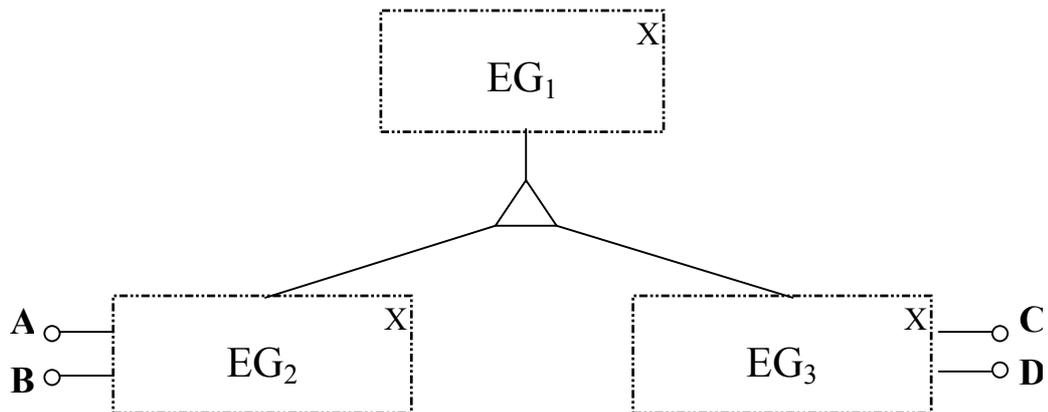
Mapping Attributi geometrici/Strati

EG₂.G₂ → S_EG₁ oppure S

EG₃.G₃ → S_EG₁ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI COMPOSIZIONE tra tipi di entità geometriche



Schema logico/fisico

Relazioni:

$EG_1(G_1: \text{POINT/LINE/POLYGON})$

$EG_2(\underline{A}, \underline{B}, G_2: \text{POINT/LINE/POLYGON})$

$EG_3(\underline{C}, \underline{D}, G_3: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure

$S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

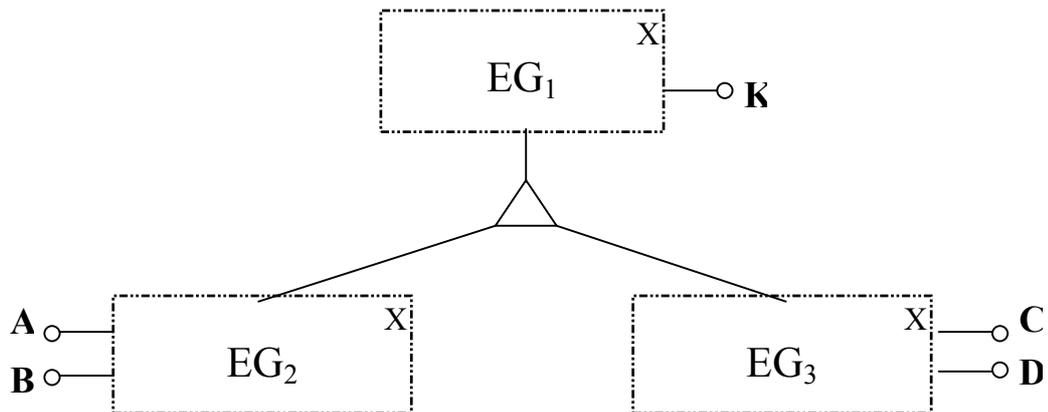
$EG_1.G_1 \rightarrow \text{Aggregate}(S_EG_1)$ oppure $\text{Aggregate}(S)$

$EG_2.G_2 \rightarrow S_EG_1$ oppure S

$EG_3.G_3 \rightarrow S_EG_1$ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI COMPOSIZIONE tra tipi di entità geometriche



Schema logico/fisico

Relazioni:

$EG_1(\underline{K}, G_1: \text{POINT/LINE/POLYGON})$

$EG_2(\underline{A}, \underline{B}, \underline{K}, G_2: \text{POINT/LINE/POLYGON})$

$EG_3(\underline{C}, \underline{D}, \underline{K}, G_3: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure

$S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow \text{Aggregate}(S_EG_1)$ oppure $\text{Aggregate}(S)$

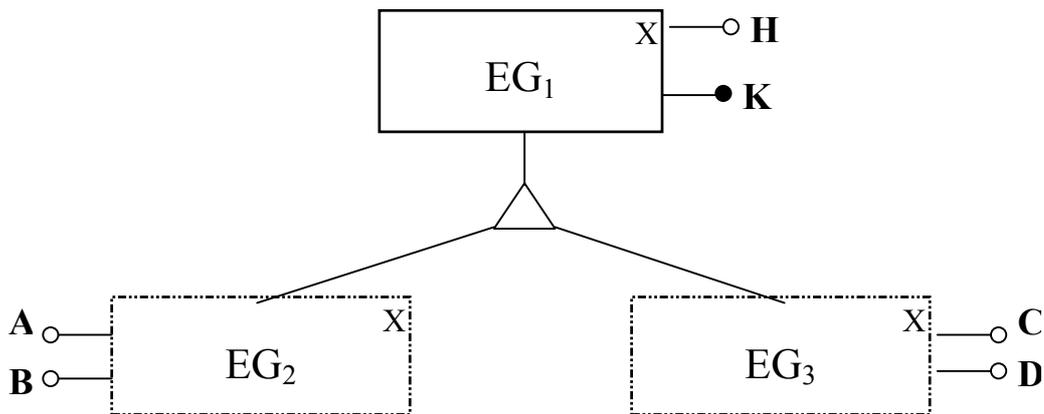
$EG_2.G_2 \rightarrow S_EG_1$ oppure S

$EG_3.G_3 \rightarrow S_EG_1$ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI COMPOSIZIONE tra entità normali con un attributo geometrico e entità geometriche.

Caso A: entità normale padre della gerarchia di composizione



Schema logico/fisico

Relazioni:

$EG_1(\underline{K}, H, G_1: \text{POINT/LINE/POLYGON})$

$EG_2(\underline{A}, \underline{B}, \underline{K}, G_2: \text{POINT/LINE/POLYGON})$

$EG_3(\underline{C}, \underline{D}, \underline{K}, G_3: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure

$S(\{\text{POINT/LINE/POLYGON}, \dots\})$

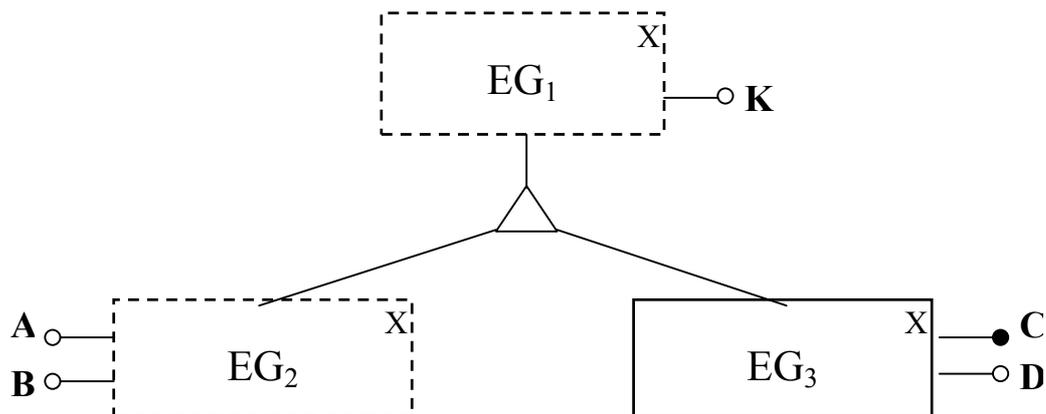
Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow \text{Aggregate}(S_EG_1)$ oppure $\text{Aggregate}(S)$

$EG_2.G_2 \rightarrow S_EG_1$ oppure S

$EG_3.G_3 \rightarrow S_EG_1$ oppure S

Caso A: entità normale figlia della gerarchia di composizione



Schema logico/fisico

Relazioni:

$EG_1(\underline{K}, G_1: \text{POINT/LINE/POLYGON})$

$EG_2(\underline{A}, \underline{B}, \underline{K}, G_2: \text{POINT/LINE/POLYGON})$

$EG_3(\underline{C}, D, K, G_3: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure

$S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

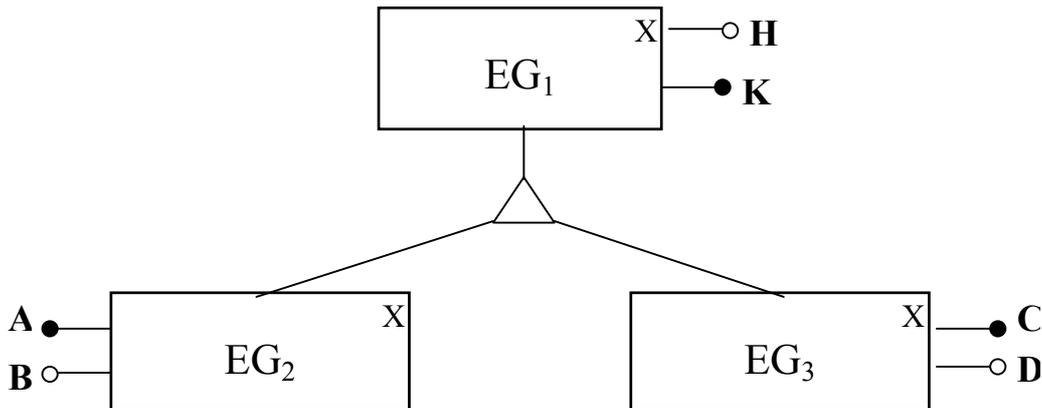
$EG_1.G \rightarrow \text{Aggregate}(S_EG_1)$ oppure $\text{Aggregate}(S)$

$EG_2.G_2 \rightarrow S_EG_1$ oppure S

$EG_3.G_3 \rightarrow S_EG_1$ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI COMPOSIZIONE tra entità normali con un attributo geometrico.



Schema logico/fisico

Relazioni:

EG₁(K, H, G₁: POINT/LINE/POLYGON)

EG₂(A, B, K, G₂: POINT/LINE/POLYGON)

EG₃(C, D, K, G₃: POINT/LINE/POLYGON)

Strati

S_EG₁({POINT/LINE/POLYGON}) oppure

S({POINT/LINE/POLYGON,...})

Mapping Attributi geometrici/Strati

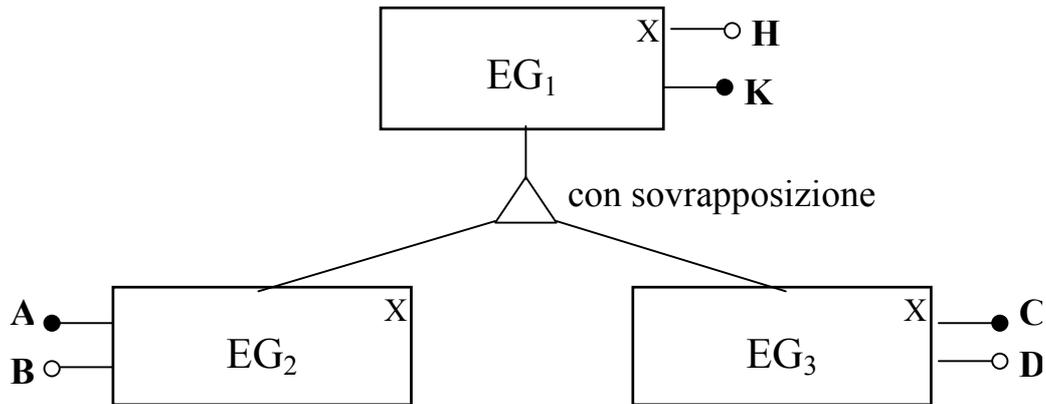
EG₁.G → Aggregate(S_EG₁) oppure Aggregate(S)

EG₂.G₂ → S_EG₁ oppure S

EG₃.G₃ → S_EG₁ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI COMPOSIZIONE tra entità normali con un attributo geometrico.



Schema logico/fisico

Relazioni:

EG₁(K, H, G₁: POINT/LINE/POLYGON)

EG₂(A, B, G₂: POINT/LINE/POLYGON)

EG₃(C, D, G₃: POINT/LINE/POLYGON)

COMPOSIZIONE_G₁_G₂(K, A)

COMPOSIZIONE_G₁_G₃(K, C)

Strati

S_EG₁({POINT/LINE/POLYGON}) oppure

S({POINT/LINE/POLYGON,...})

Mapping Attributi geometrici/Strati

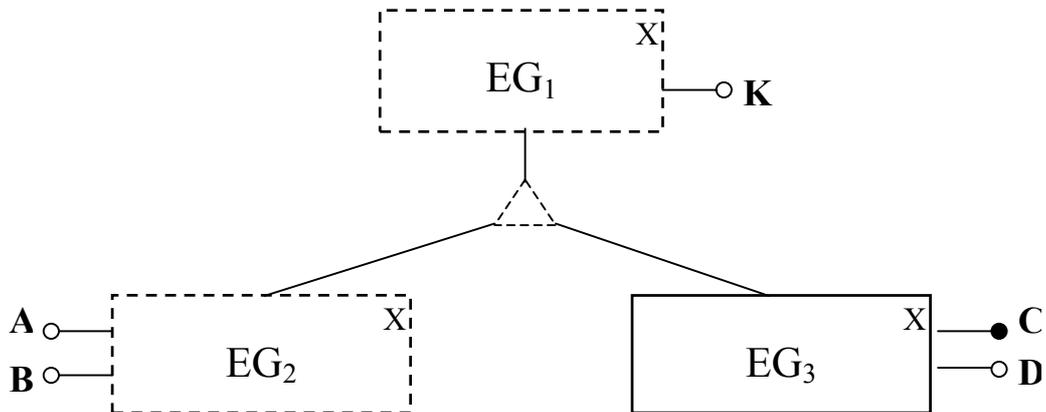
EG₁.G → Aggregate(S_EG₁) oppure Aggregate(S)

EG₂.G₂ → S_EG₁ oppure S

EG₃.G₃ → S_EG₁ oppure S

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

GERARCHIE DI CONTENIMENTO tra tipi di entità geometriche o normali con attributo geometrico.



Schema logico/fisico

Relazioni:

$EG_1(\underline{K}, G_1: \text{POINT/LINE/POLYGON})$

$EG_2(\underline{A}, \underline{B}, \underline{K}, G_2: \text{POINT/LINE/POLYGON})$

$EG_3(\underline{C}, \underline{D}, K, G_3: \text{POINT/LINE/POLYGON})$

Strati

$S_EG_1(\{\text{POINT/LINE/POLYGON}\})$ oppure

$S(\{\text{POINT/LINE/POLYGON}, \dots\})$

Mapping Attributi geometrici/Strati

$EG_1.G \rightarrow S_EG_1$ oppure S

$EG_2.G_2 \rightarrow S_EG_1$ oppure S

$EG_3.G_3 \rightarrow S_EG_1$ oppure S

Si noti che:

- per le entità geometriche la relazione di contenimento determina le istanze delle entità figlie, quindi va memorizzata esplicitamente attraverso l'inserimento della chiave del padre nei figli.

Le regole di corrispondenza tra GEO-ER e il modello geo-relazionale

VINCOLO DI INTERSEZIONE NON NULLA tra tipi di entità geometriche o normali con attributo geometrico.

Esprime solo un vincolo geometrico. E' quindi consigliabile rappresentare nello stesso strato i valori degli attributi geometrici (o istanze di entità geometriche) coinvolti nel vincolo.

Si noti in particolare che un'intersezione non nulla tra punti e linee equivale ad una gerarchia di contenimento.

Esercizio

Scrivere lo schema concettuale utilizzando il modello GEO-ER per le seguenti informazioni.

Si vogliono memorizzare i dati che descrivono l'uso del suolo di una porzione di territorio. Si devono individuare i poligoni che rappresentano le aree urbanizzate, le aree non urbanizzate, i corpi idrici, le cave e le discariche.

Le aree urbanizzate sono suddivise in: aree verdi, aree residenziali, aree produttive, aree miste, aree in trasformazione, servizi e aree di infrastrutture; queste ultime sono suddivise in aree stradali o autostradali, stazioni ferroviarie e scali. I servizi sono i porti, gli aeroporti, dogana, università, ospedale, cimitero, chiesa, carcere, rifugio alpino, terme.

Le aree non urbanizzate sono suddivise in: aree rocciose, boschi, aree sterili, aree non coltivate, e aree agricole; queste ultime si suddividono in: aree legnose agrarie, aree a prato o pascolo e aree a seminativo.

Le cave sono individuate da un codice, un nome, da un tipo (cava, miniera, torbiera) da uno stato (in esercizio, non in esercizio).

Le discariche sono individuate da un codice, un nome, da uno stato (in esercizio, non in esercizio).