

Corso
di
Basi di Dati Spaziali

Introduzione

Angelo Montanari
Donatella Gubiani

Sistemi Informativi Geografici e Basi di Dati Spaziali

- Qual è il legame tra Sistemi Informativi Geografici/Territoriali (GIS) e Basi di Dati Spaziali?
- Analogia col legame tra Sistemi Informativi e Basi di Dati

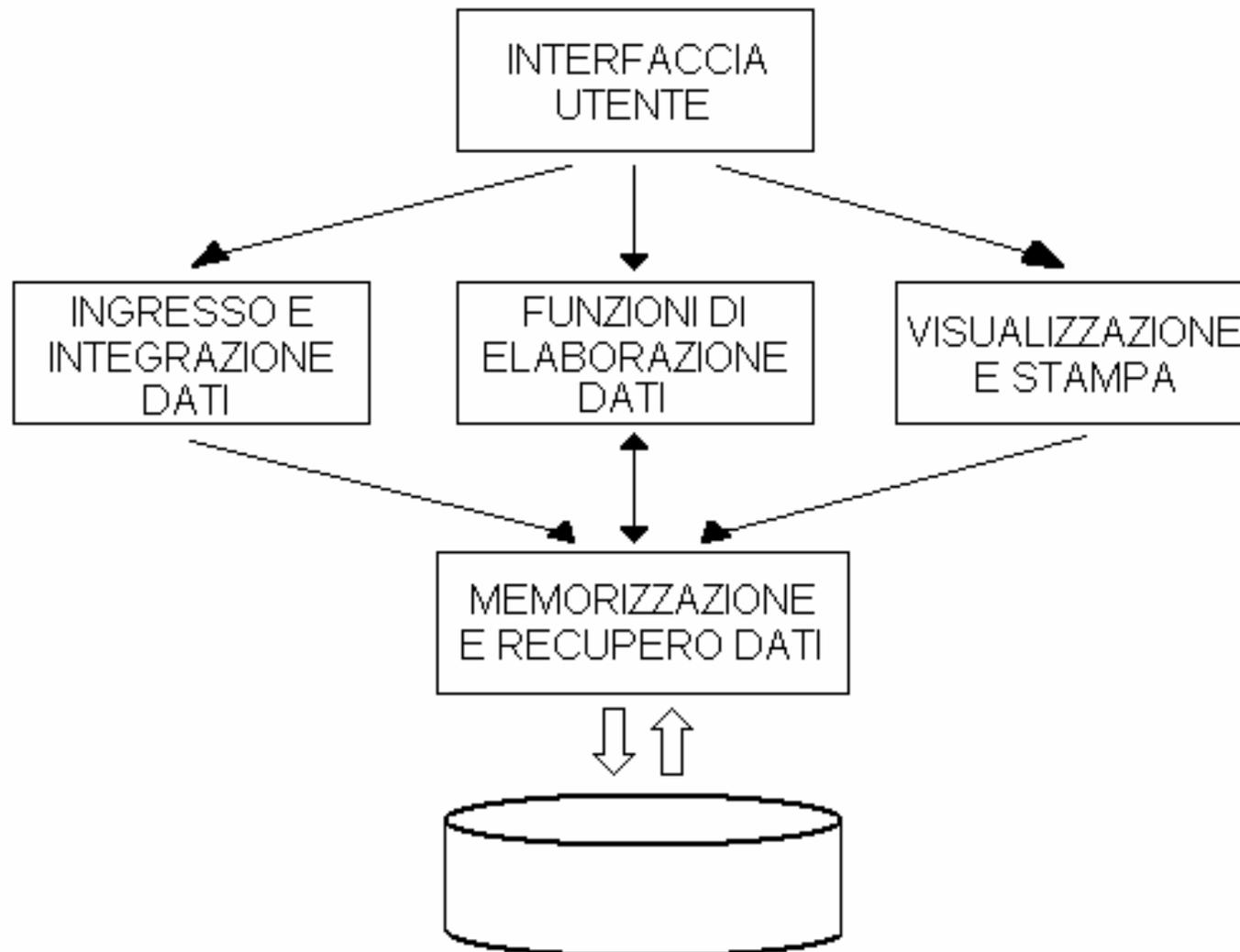
Sistema Informativo

- Componente (sottosistema) di una organizzazione che gestisce le informazioni di interesse:
 - raccolta/acquisizione
 - archiviazione/conservazione
 - elaborazione, trasformazione, produzione
 - distribuzione, comunicazione, scambio

Sistemi Informativi Geografici (GIS)

- Sistemi automatizzati per memorizzare, analizzare e manipolare dati geografici
- Diversi (altri) sistemi:
 - Computer-Aided Design (CAD)
 - Computer Cartography System
 - Remote Sensing System

Architettura GIS



GIS e cartografia

- In passato l'informazione geografica era rappresentata e conservata attraverso carte geografiche
- Con l'avvento dell'informatica, sono stati sviluppati sistemi informativi geografici automatizzati
 - primi tentativi anni '50
 - obiettivo: semplificazione dei processi di produzione, utilizzo e mantenimento delle mappe

Convenzioni

- **SISTEMA DI COORDINATE**
 - localizzazione degli oggetti sulla superficie terrestre
 - latitudine e longitudine correlate alla quota sul livello del mare
- **PROIEZIONI CARTOGRAFICHE**
 - la superficie terrestre è sferica e deve essere portata su una mappa bidimensionale
 - distorsioni

Concetto di scala

- **Cartografia:**
 - proporzione fra le distanze misurate sulla carta e le distanze reali
 - indica implicitamente la precisione e il grado massimo di dettaglio raggiungibile
- **GIS:**
 - rappresenta la precisione dei dati spaziali rispetto al reale
 - caratterizza il repertorio degli oggetti e la loro simbologia

Caratteristiche della Cartografia

- La precisione della carta fa parte della meta-informazione ad essa associata
- Tutti i simboli hanno la medesima precisione
- Le relazioni spaziali fra gli oggetti sono derivabili dall'analisi della posizione relativa dei simboli sulla carta
- Non può essere aggiornata, deve essere completamente rigenerata

Caratteristiche dei GIS

- Rappresentano in modo esplicito le coordinate reali
- La meta-informazione sulla precisione metrica può essere descritta attraverso il concetto di granularità
- Le relazioni spaziali fra gli oggetti possono essere specificate esplicitamente
- Possono essere aggiornati con continuità

Componenti di un oggetto geografico

- Attributo geometrico/spaziale (o estensione spaziale), che specifica posizione, forma, orientazione e dimensione dell'oggetto nello spazio 2D o 3D
- attributi non spaziali, detti attributi tematici o descrittivi (ad esempio, il nome di una particella)

Oggetto del corso

- Concetti, tecniche e algoritmi utilizzati per la gestione di oggetti geografici/spaziali in basi di dati spaziali
- Basi di dati spaziali vs. basi di immagini (le basi di immagini supportano la ricerca di scene che contengono certi oggetti o pattern, ignorando la loro posizione, orientazione e dimensione)

Problematiche affrontate

- Come rappresentare i dati?
- Quali operazioni sono necessarie per recuperare e manipolare tali dati?
- Come gestire dati in memoria secondaria?

Caratteristiche fondamentali

- **Grande volume** di dati (la quantità di immagini registrate giornalmente da un satellite è dell'ordine dei terabyte)
- **Struttura** intrinsecamente **complessa** dei dati

Origine dei dati - 1

- Raccolti direttamente sul campo (col supporto del GPS satellitare, immagini via satellite, fotogrammetria, ..)
- Possibili problemi:
 - delimitazione degli oggetti geografici
 - associazione degli attributi tematici agli oggetti spaziali

Origine dei dati - 2

- Ricavati da mappe esistenti
- Possibili problemi (qualora si debbano integrare sorgenti diverse):
 - eterogeneità delle scale e dei sistemi di coordinate
 - diversa qualità dei dati

Vocabolario delle applicazioni - 1

- **Tema:** informazione geospaziale relativa ad un particolare aspetto. Ha uno schema e delle istanze. Esempi: fiumi, città, stati.
- Un tema è espresso attraverso una mappa.
- Progettazione: costruzione di uno schema concettuale per ogni tema di interesse

Vocabolario delle applicazioni - 2

- **Oggetto geografico/spaziale:**
corrisponde ad un'entità del mondo reale ed ha 2 componenti:
 - una *descrizione* (insieme di attributi descrittivi, quali popolazione e nome di una città)
 - una *componente spaziale*, che comprende sia gli aspetti geometrici sia quelli topologici

Vocabolario delle applicazioni - 3

- **Aspetti geometrici**

localizzazione sul sottostante spazio geografico, forma, ..

- **Aspetti topologici**

relazioni spaziali tra oggetti, quali, ad esempio, l'adiacenza o l'inclusione

Vocabolario delle applicazioni - 4

- La componente spaziale di un oggetto geografico è detta **oggetto spaziale**

Può essere presa in considerazione separatamente, ad esempio, quando è condivisa da più entità geografiche (è questo il caso del confine tra due stati)

P.S. In alcune applicazioni vi è una terza componente, la *componente temporale*.

Oggetti atomici e complessi

- Oggetti geografici **atomici**
- Oggetti geografici **complessi** composti da altri oggetti geografici, atomici o complessi

Esempio. Nel *tema* "unità amministrativa americana", l'oggetto complesso "stato della California" è composto dagli oggetti atomici "contee della California"

Temi e oggetti geografici

- Un tema è una collezione di oggetti geografici **omogenei**, ossia aventi la stessa struttura/dello stesso tipo (particolare astrazione dello spazio che fa riferimento allo stesso unico tipo di oggetti)

Esempio. Il *tema* delle città, ognuna descritta attraverso gli stessi attributi (*schema* del tema), ad esempio, nome, popolazione e geometria.

Modello geografico

- Ogni tema va rappresentato in termini del modello logico dei dati sottostante (**modello geografico**)

Tale modello consente di descrivere e manipolare i temi e i loro oggetti in un DBMS

Modelli dei dati spaziali

- L'attributo spaziale di un oggetto geografico non corrisponde ad alcun tipo di dato standard
- Per modellare geometria e topologia degli oggetti geografici occorrono strumenti potenti a livello di temi e di oggetti (**modelli dei dati spaziali**)

Tipi di dato di base

- Punto (oggetti di dimensione 0)
 - Linea (dimensione 1)
 - Regione/poligono (dimensione 2)
- Esempio.** Linee per i fiumi, regioni per le città

Inoltre,

- Altezza di un punto, ad esempio di un edificio (dimensione 2,5)
- Volume (dimensione 3)
- Spazio-tempo (dimensione 4)

Manipolazione dei dati geospaziali

- Operazioni semplici (“algebra dei temi”): prendono in ingresso uno o più temi e restituiscono un tema
- Operazioni complesse (operazioni che usano una metrica, operazioni topologiche, ..)
- Operazioni tipiche dei GIS

Operazioni semplici - 1

- **Proiezione**

Π : tema $X \{A_1, \dots, A_n\} \rightarrow$ tema

Input: un tema e un sottoinsieme dei suoi attributi (incluso, esplicitamente o implicitamente, quello spaziale)

Output: un nuovo tema con attributi descrittivi A_1, \dots, A_n e parte spaziale uguale a quella del tema iniziale

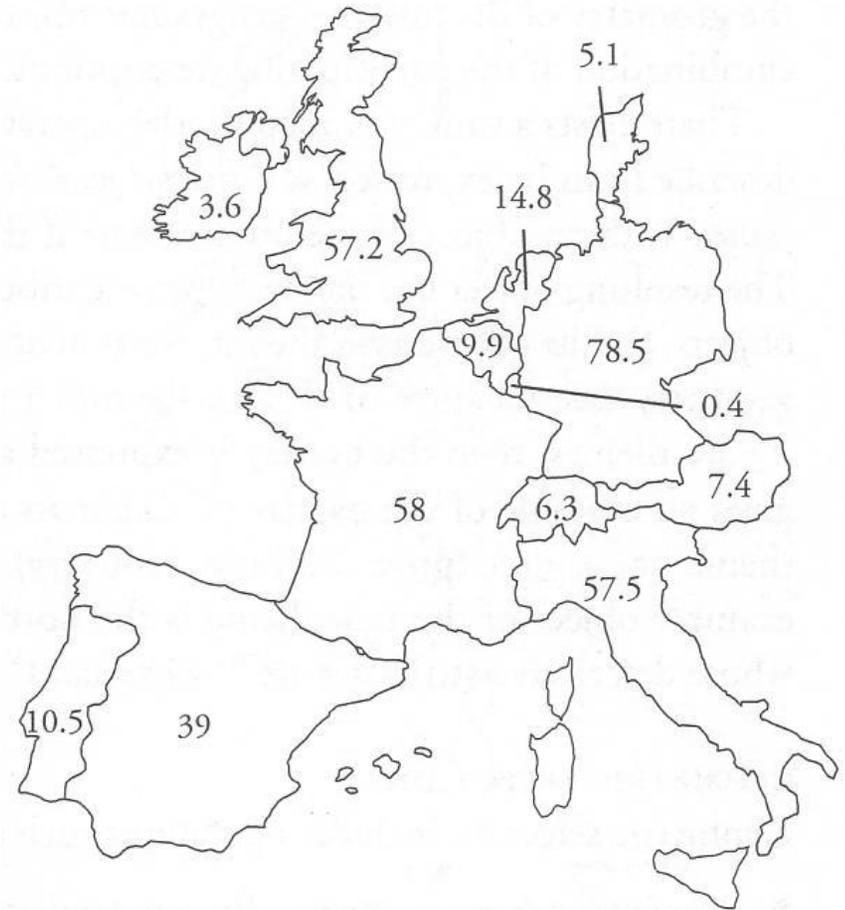
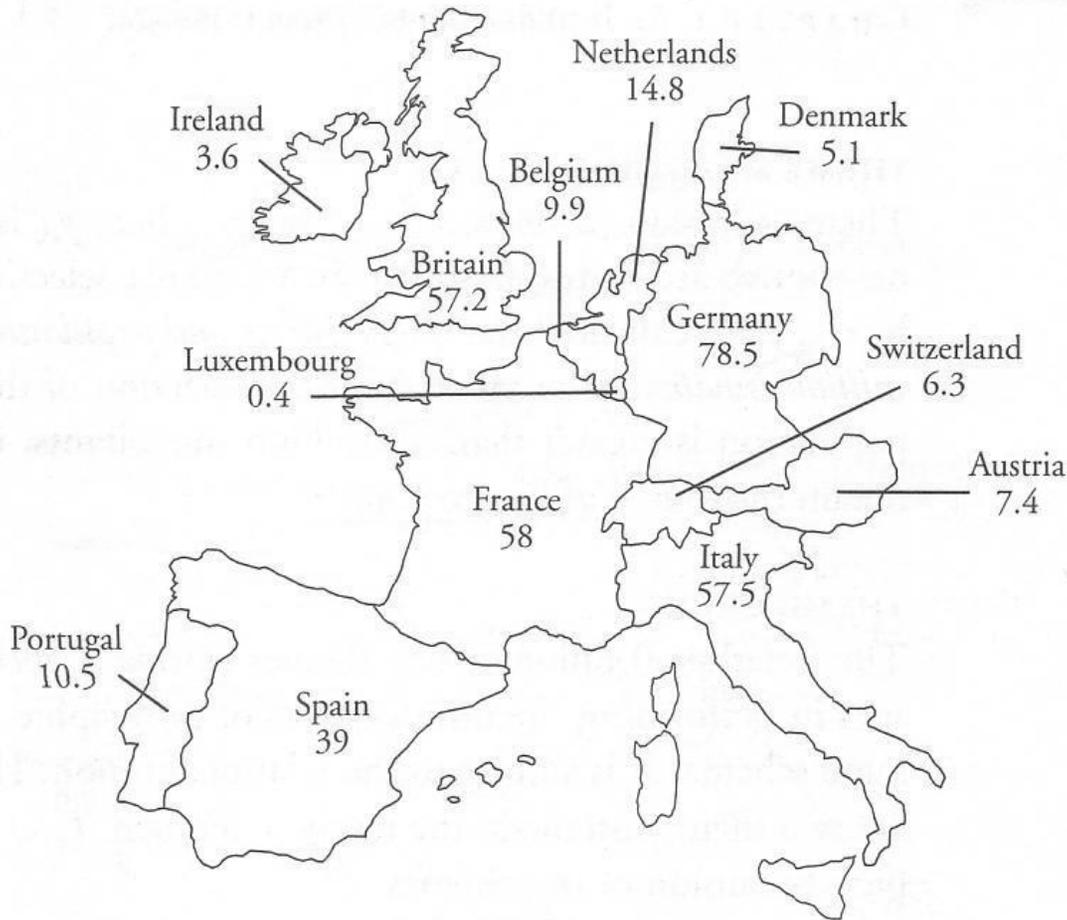
Operazioni semplici - 2

Esempio. Data un'istanza T del tema stato,

$$\Pi_{\text{Popolazione,geo}}(T)$$

restituisce la parte spaziale corredata della sola informazione relativa alla popolazione

Operazioni semplici – 2 a



Operazioni semplici - 3

- **Selezione**

σ : tema X $p_{Ai} \rightarrow$ tema

Input: un tema e un predicato che coinvolge i soli attributi descrittivi

Output: un nuovo tema che contiene tutti e soli gli oggetti che soddisfano la condizione

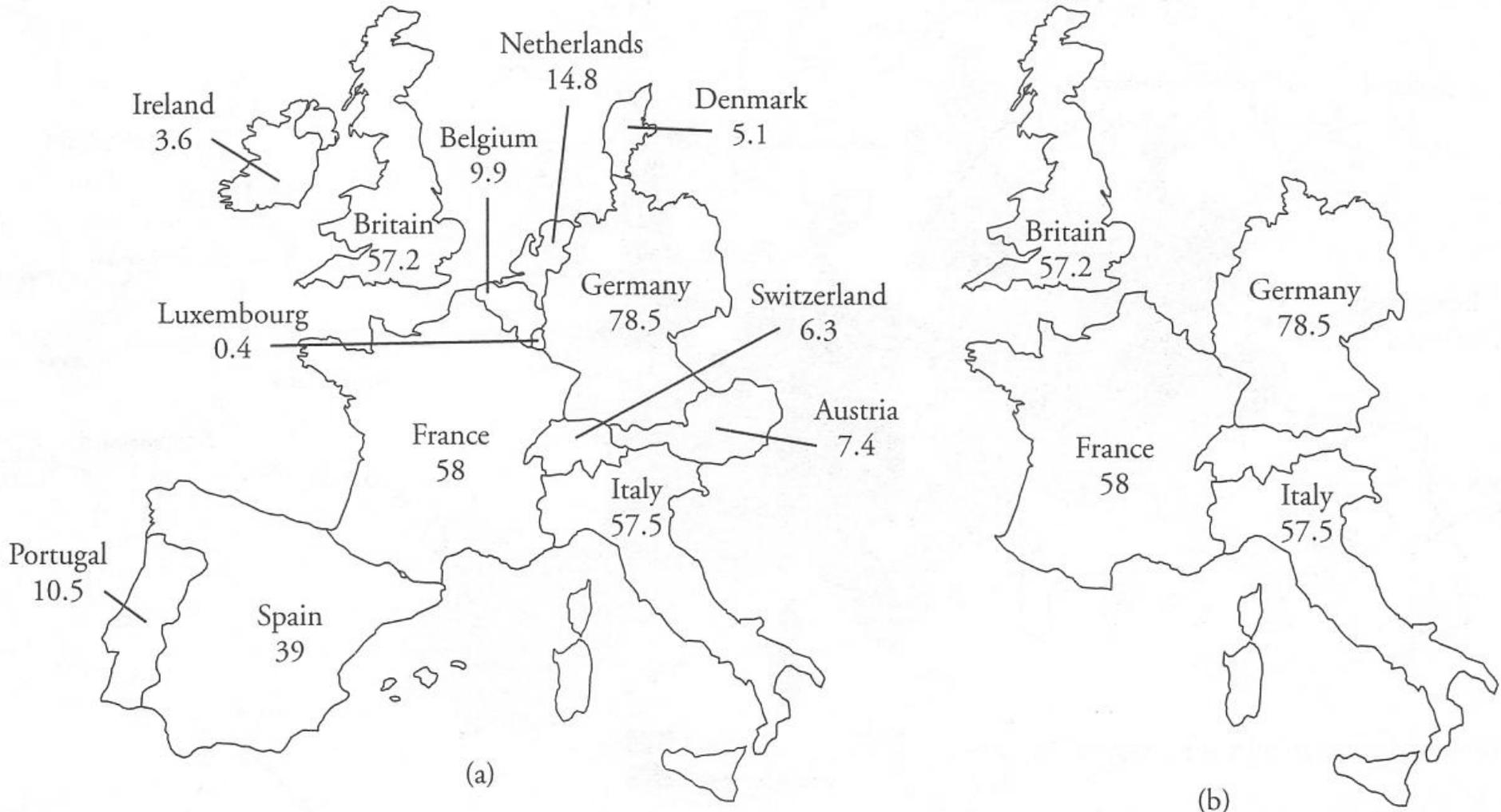
Operazioni semplici - 4

Esempio. Data un'istanza T del tema stato,

$$\sigma_{\text{Popolazione} \geq 50.000.000}(T)$$

restituisce tutti e soli gli stati, con i loro attributi descrittivi e l'attributo spaziale, con una popolazione di 50.000.000 di abitanti o più

Operazioni semplici – 4 a



Operazioni semplici - 5

Unione

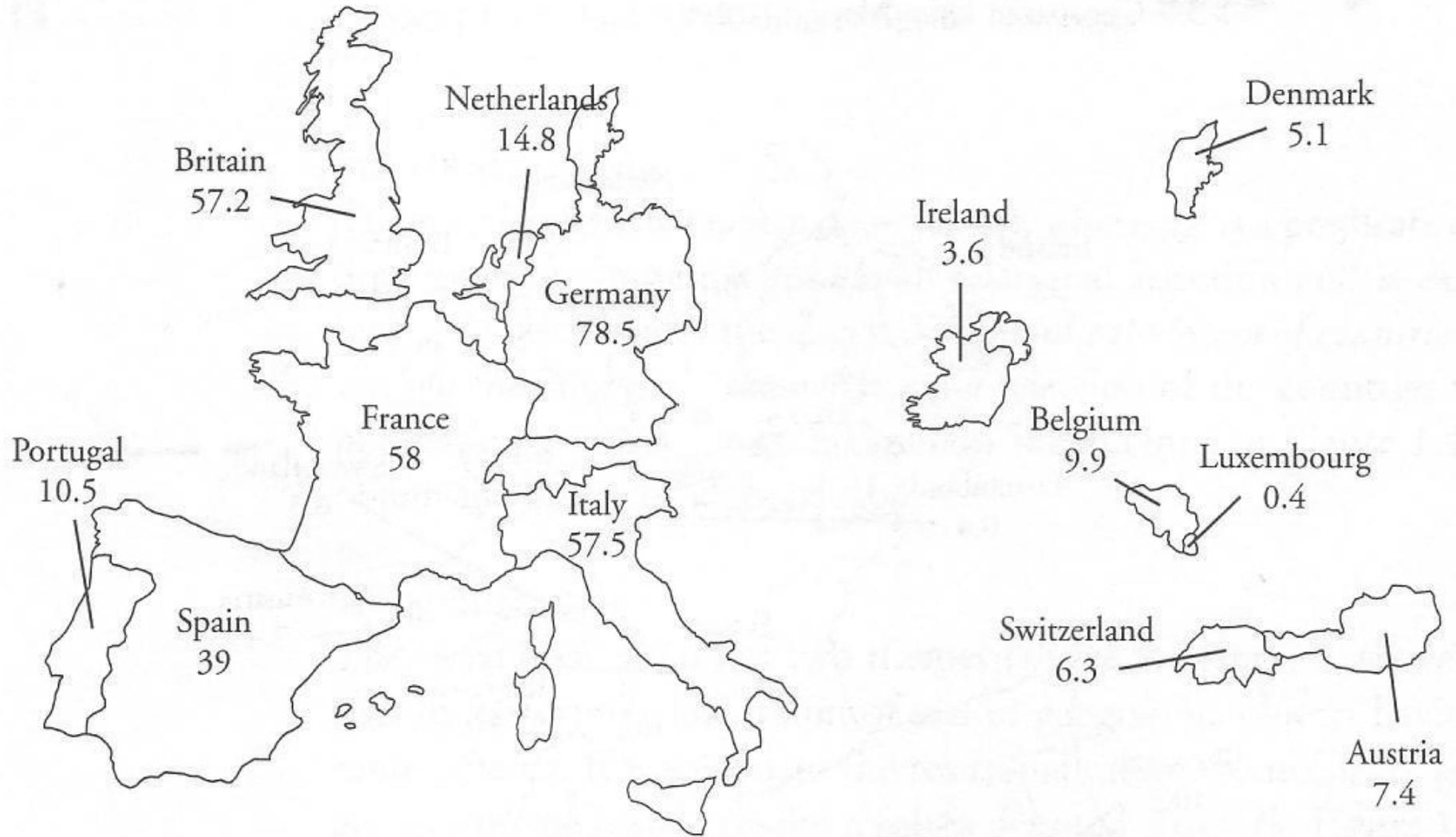
\cup : tema X tema \rightarrow tema

Input: due temi aventi lo stesso schema

Output: un tema contenente gli oggetti appartenenti ad almeno uno dei due temi ricevuti in input

Esempio. Unione degli stati europei con meno di 10.000.000 di abitanti con quelli aventi 10.000.000 o più abitanti

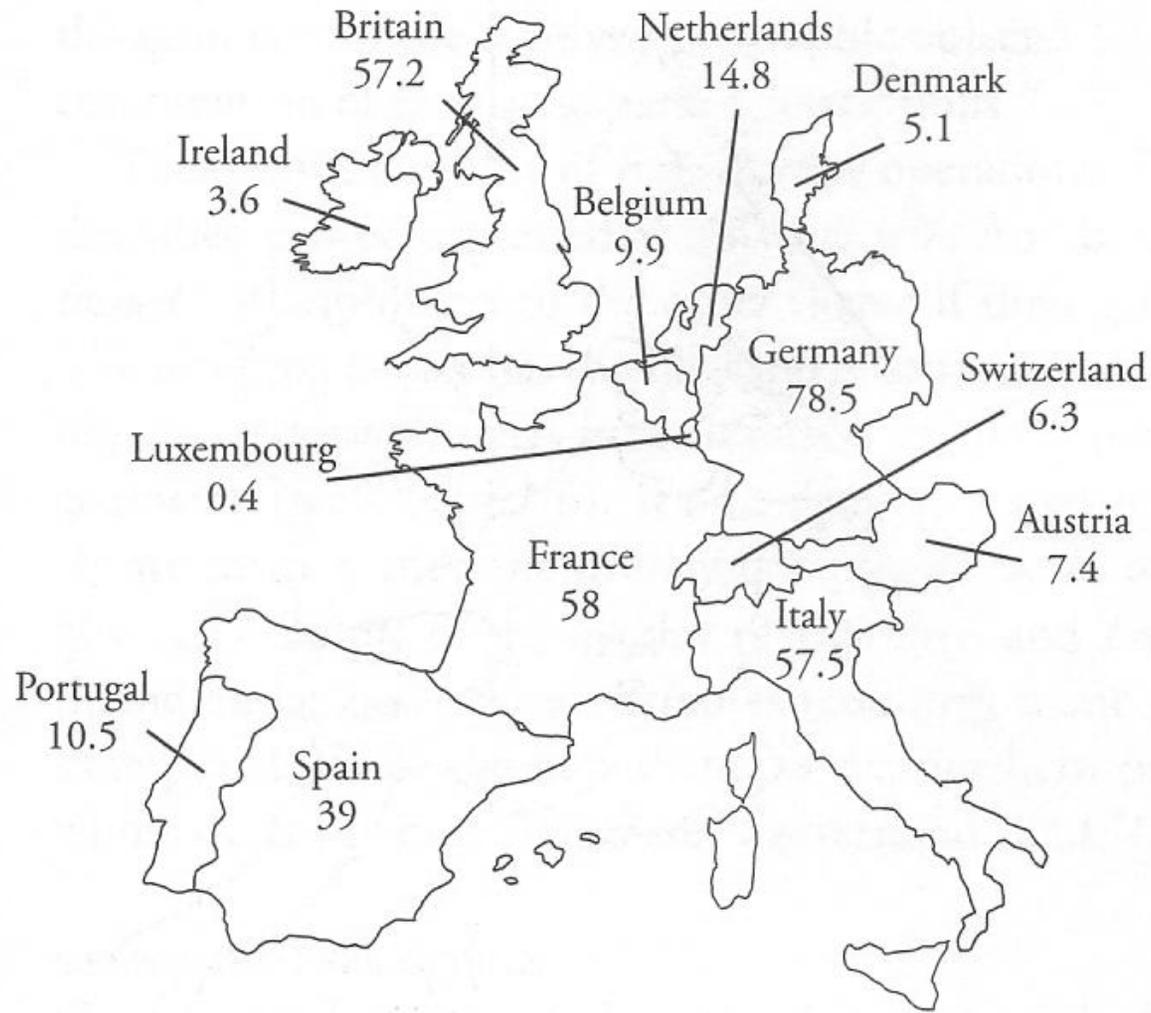
Operazioni semplici – 5 a



T_1

T_2

Operazioni semplici – 5 b



$$T_1 \cup T_2$$

Operazioni semplici - 6

Overlay (copertura)

tema X tema → tema

- E' un'operazione tipica delle applicazioni GIS, di cui esistono più varianti
- Genera un nuovo tema, a partire da temi che si sovrappongono, che contiene nuovi oggetti geografici.

Operazioni semplici - 7

Ad esempio, il **join spaziale**

- La *geometria* degli oggetti del tema risultante è calcolata intersecando le geometrie degli oggetti coinvolti
- La *descrizione* degli oggetti del tema risultante è una combinazione delle descrizioni degli oggetti coinvolti nell'operazione

Operazioni semplici - 8

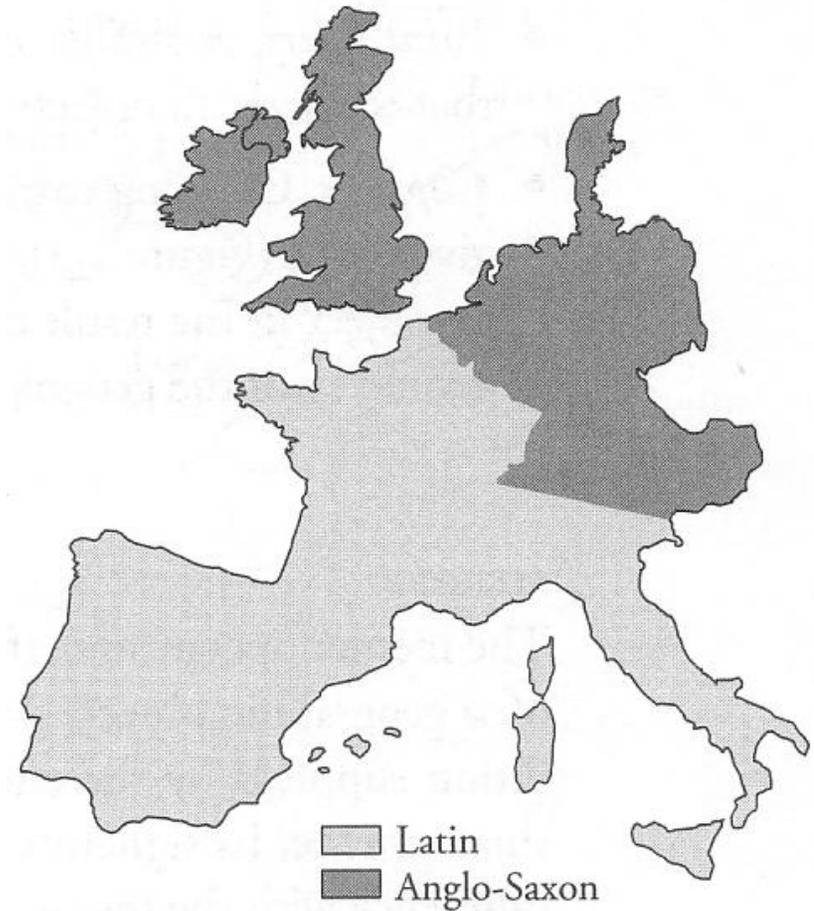
Esempio. Data un'istanza T_1 del tema stato e un'istanza T_2 del tema linguaggio (comprendente due oggetti: di radice latina e di radice anglosassone), il join spaziale può essere utilizzato per individuare, all'interno di ciascuno stato, le diverse aree linguistiche.

Ad esempio, al nord della Svizzera verrà associata la radice anglosassone.

Operazioni semplici – 8 a



T₁



T₂

Operazioni semplici – 8 b

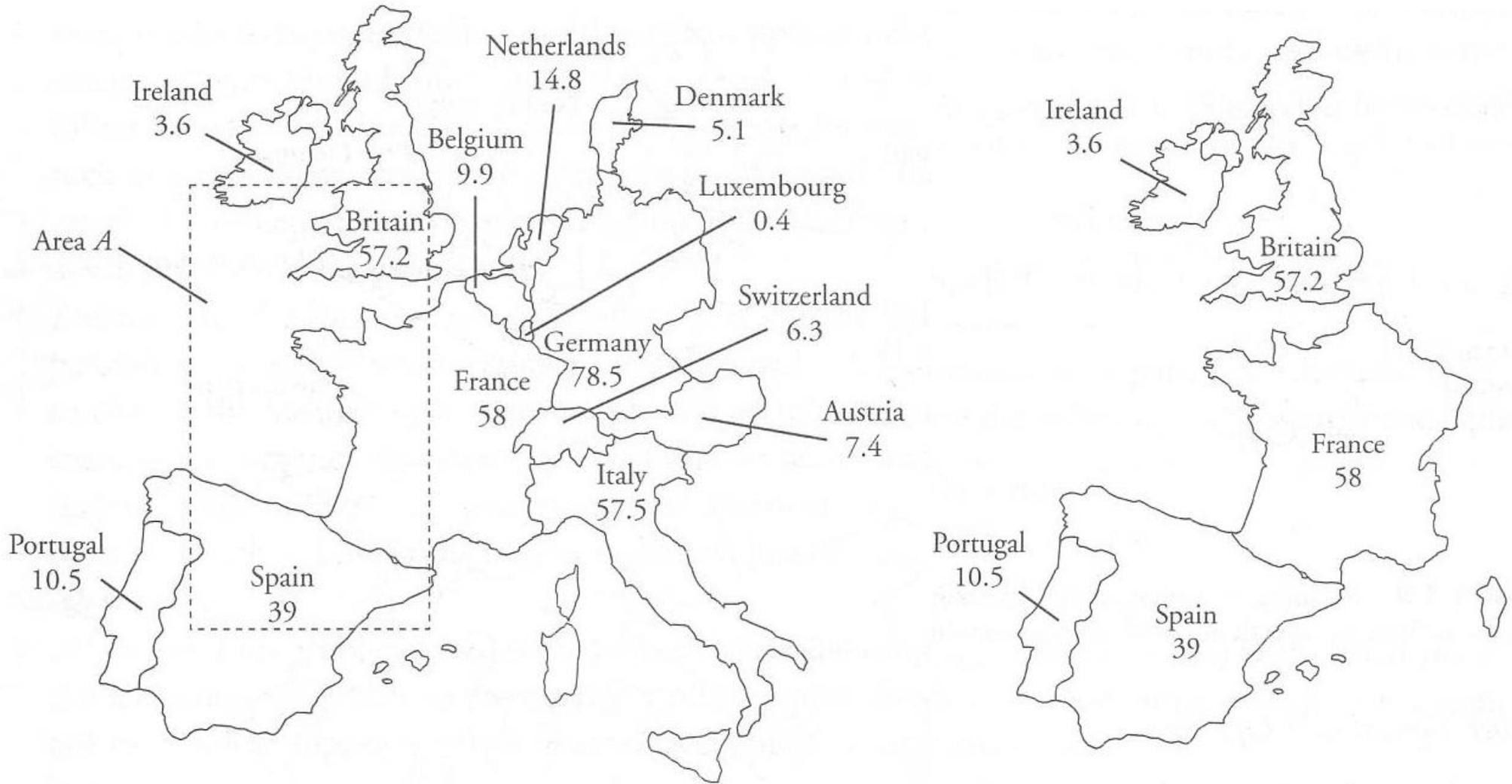


Overlay (T_1, T_2)

Operazioni semplici - 9

- **Selezione geometrica**
comprende più operazioni quali
 - **Windowing** (window query): dato un tema restituisce un tema che contiene tutti e soli gli oggetti del tema iniziale che hanno (la cui geometria ha) intersezione non vuota con una data area/finestra, usualmente di forma rettangolare

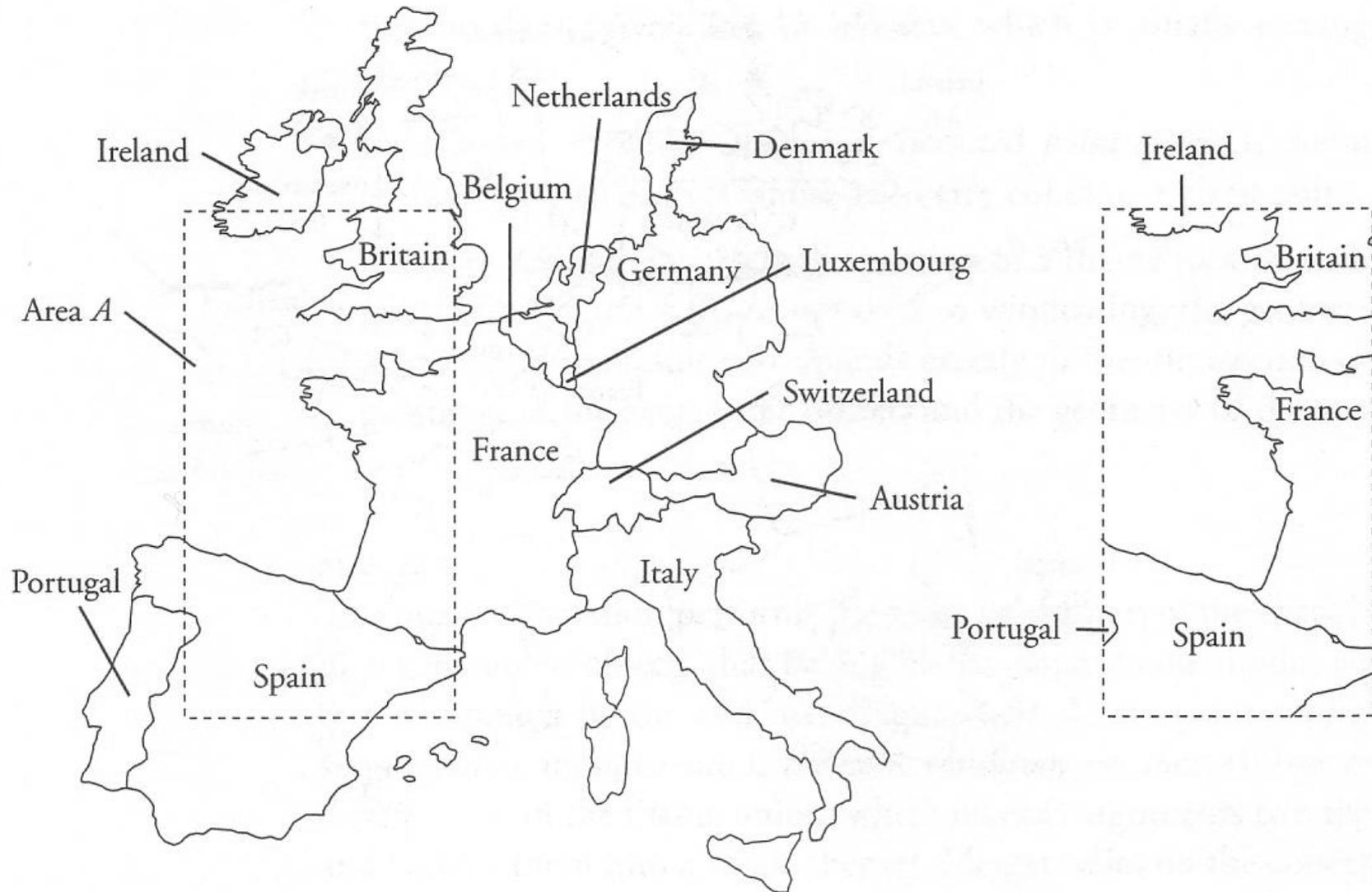
Operazioni semplici – 9 a



Operazioni semplici - 10

- **Interrogazione puntuale** (point query): dato un tema restituisce un tema che contiene tutti e soli gli oggetti del tema iniziale che contengono un dato punto
- **Clipping**: estrae la porzione del tema iniziale localizzata all'interno di una data area (intersezione della geometria degli oggetti di input e della geometria dell'area)

Operazioni semplici – 10 a



Operazioni semplici - 11

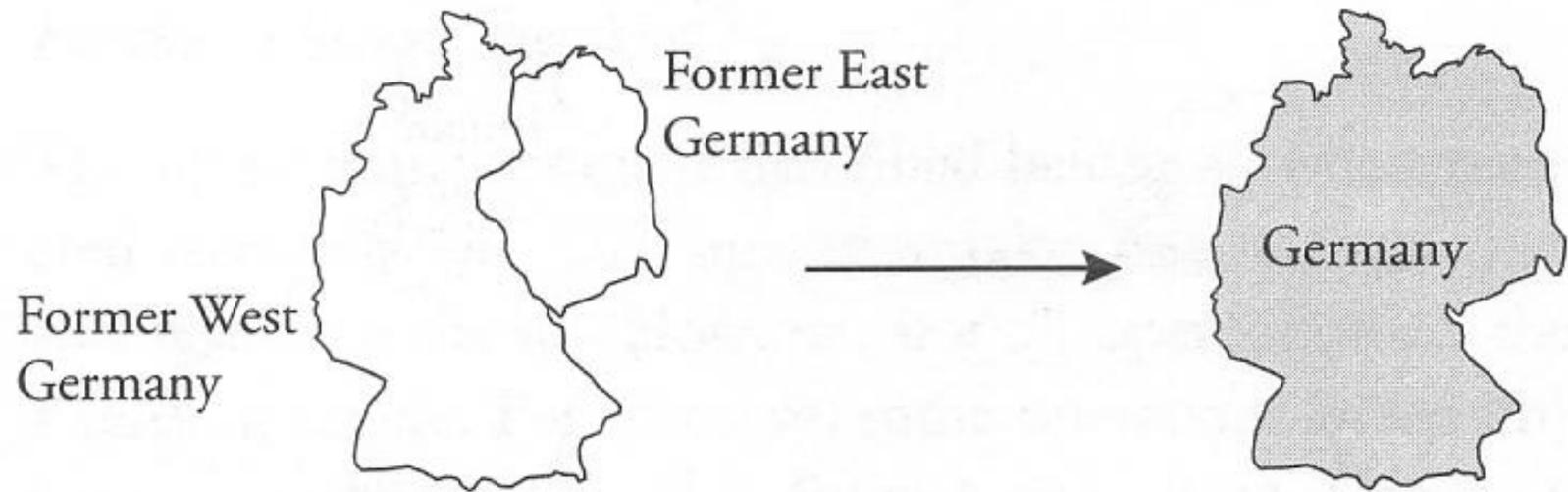
- **Merger** (Fusione)

tema X condizione → tema

unione geometrica della componente spaziale di n oggetti di un dato tema in base a una condizione (legata alla nozione di **aggregazione di oggetti**)

Esempio. Individuazione del territorio di una regione a partire da quelli delle sue province.

Operazioni semplici – 11 a



Operazioni complesse

Non tutte le operazioni restituiscono un tema. Alcune operazioni sono più complesse.

Alcuni esempi:

- operazioni che usano una **metrica** (che consentono, ad esempio, di determinare la distanza tra due città)
- **operazioni topologiche** (che consentono, ad esempio, di determinare gli stati che confinano con la Svizzera)

Operazioni tipiche dei GIS

Esistono diverse altre operazioni tipiche dei GIS utilizzate, ad esempio, per creare temi/mappe o per l'analisi di dati geospaziali.

Esse comprendono le operazioni di interpolazione/estrapolazione (ad esempio, in presenza di un numero limitato di dati ottenuto da un campionamento), posizionamento, allocazione, mix di posizionamento e allocazione.

DBMS come componente di un GIS - 1

Diversi approcci

- Uso di un **DBMS relazionale standard** (inadeguato)

Esempio. Il tema stato può essere rappresentato da una tabella **stato** contenente gli attributi descrittivi più un attributo (codice) che ne specifica la frontiera. Dato che uno stato può essere costituito da più parti separate (è questo il caso dell'Italia), una tabella **frontiera** specifica, per ogni parte, i contorni che la definiscono. Una tabella **contorno** associa ad ogni contorno (una poligonale) un insieme di punti (i vertici della poligonale), specificando l'ordine di tali punti nella poligonale. Infine, una tabella **punto** fornisce le coordinate di ciascun punto.

DBMS come componente di un GIS - 2

Limiti di tale soluzione:

- violazione del principio di indipendenza dei dati (per interrogare i temi occorre conoscere la struttura dei dati spaziali)
- basse prestazioni (vedi esempio)
- poco amichevole (non è user-friendly)
- difficoltà/impossibilità di esprimere computazioni geometriche (test di adiacenza, window query, point query)

DBMS come componente di un GIS - 3

- **Accoppiamento debole**

gestione separata dei dati descrittivi, usualmente con un DBMS relazionale, e dei dati geometrici (è la soluzione adottata da sistemi GIS quali ArcInfo-ESRI, MGE, TiGRis-Intergraph)

Limiti di tale soluzione:

- coesistenza di modelli dei dati eterogenei
- perdita di funzionalità di base di un DBMS (interrogazione, ottimizzazione, ripristino)

DBMS come componente di un GIS - 4

- **Approccio integrato**

sfrutta l'estendibilità dei DBMS (aggiunta di nuovi tipi di dato e di nuove operazioni ad un DBMS relazionale).

E' la soluzione adottata da Oracle e Postgres.

DBMS come componente di un GIS - 5

Principali estensioni:

- il linguaggio di interrogazione SQL viene arricchito per poter manipolare sia dati tradizionali (descrittivi) sia dati spaziali. Vengono inoltre introdotti nuovi tipi di dato spaziale (punto, linea, regione)
- diverse altre funzionalità dei DBMS (ad esempio, l'ottimizzazione delle interrogazioni) vengono adattate per consentire una gestione efficiente dei dati geospaziali

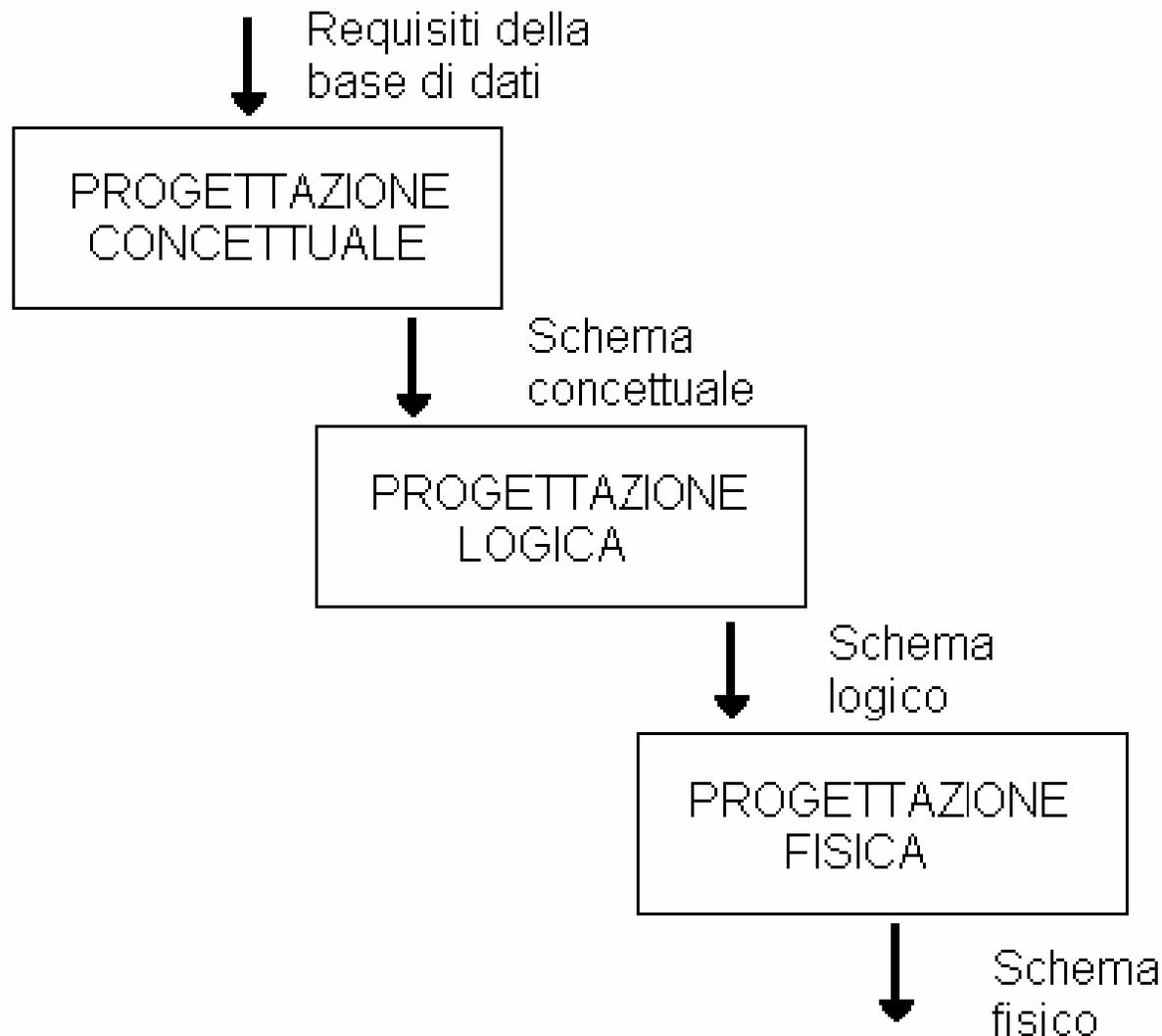
Linguaggi per basi di dati spaziali - 1

- Analogamente a quanto accade nelle basi di dati tradizionali, esistono diversi tipi di linguaggi:
 - linguaggi di definizione dei dati (DDL)
 - linguaggi di manipolazione dei dati (DML):
 - linguaggi di aggiornamento
 - linguaggi di interrogazione

Linguaggi per basi di dati spaziali - 2

- I linguaggio tradizionali non permettono di trattare informazioni spaziali → estensioni
- GEO-ALGEBRA: estensione spaziale dell'algebra relazionale
- Estensioni spaziali del linguaggio SQL (Oracle Spatial, PostgreSQL)

Progettazione di basi di dati spaziali -1



Progettazione di basi di dati spaziali -2

- La differenza sostanziale rispetto alla progettazione delle basi di dati tradizionali sta nei modelli utilizzati
 - estensioni dei modelli concettuali tradizionali
 - estensioni dei modelli logici tradizionali
 - strutture fisiche adeguate al trattamento di dati spaziali

Strutture di indicizzazione

- Un accesso efficiente ai dati è essenziale anche per le basi di dati spaziali
- Le strutture dati utilizzate dai DBMS relazionali (B-alberi, B+-alberi) non sono appropriate
- Occorrono nuove strutture dati

Bibliografia

- Spatial Databases (with application to GIS), P. Rigaux, M. Scholl, A. Voiosard, Morgan Kaufmann Pub., 2002
- Spatial Databases - A Tour, S. Shekhar, S. Chawla, Prentice Hall, 2003
- GIS- A Computing Perspective, M.F. Worboys, Taylor & Francis, 2001