

Dati Geografici Vettoriali

Claudio Rocchini
Istituto Geografico Militare
2018



Premessa: dati GIS Vettoriali

- Gli oggetti del mondo, giudicati importanti del punto di vista geografico ad una certa «scala», vengono rappresentati da geometrie puntuali, linee spezzate, poligoni (con buchi).
- Alle geometrie sono assegnati dei valori alfanumerici che le descrivono.
- Gli oggetti non sono alla rinfusa, ma ordinati in «classi» di oggetti (es. strade, laghi, edifici, boschi, etc.).
- Tipicamente gli oggetti di una classe sono tutti dello stesso tipo geometrico ed hanno la stessa sequenza di attributi (dello stesso tipo di dato: numeri e parole)

Dato vettoriale

- Basato sulla descrizione matematica degli oggetti che formano il mondo reale.
- Concettualmente più complesso del raster (più difficile da gestire).
- Scomposizione degli oggetti geografici in elementi geometrici con attributi associati: punti, linee, aree, (testi ?)
- Presenza di elementi topologici.

Natura del Dato Vettoriale

- I dati sono rappresentati da insiemi di coordinate
- Gli oggetti sono rappresentati da primitive geometriche:
 - Punti (ovvero punti + orientamento)
 - Curve (ovvero polilinee o spezzate)
 - Aree (ovvero poligono con o senza buchi)
- I testi di riferimento si fermano qui, ma noi (IGM) includiamo anche i testi come 4° primitiva

Confronto Raster-Vettoriale

- Meriti Raster
 - Modello di dati Semplice
 - Tecnologia Economica
 - Dati Facili da Acquisire
 - Dati Facili da Processare
- Demeriti Raster
 - Non topologia
 - Attributi limitati
 - Poca compattezza
 - Bassa qualità di visualizzazione
- Meriti Vettoriale
 - Struttura compatta
 - Processi Topologici
 - Alta qualità cartografica
 - Attributi sofisticati
- Demeriti Vettoriale
 - Modello complesso
 - Difficile op. Overlay
 - Dati costosi da acquisire
 - Tecnologia (software) costosa

Tipi di Geometrie

Geometrie Per la Cartografia

- Gli oggetti sono rappresentati da primitive geometriche, che si distinguono per alcune caratteristiche:
- Tipi fondamentali
 - Dimensione 0 : punti
 - Dimensione 1 : linee
 - Dimensione 2 : aree
 - *(Dimensione 3 : volumi, in un futuro remoto)*
 - Testi (Scelta filosofica)
- Coordinate
 - 2D
 - 2D e mezzo
 - 3D
- Altre caratteristiche: Oggetti multipli



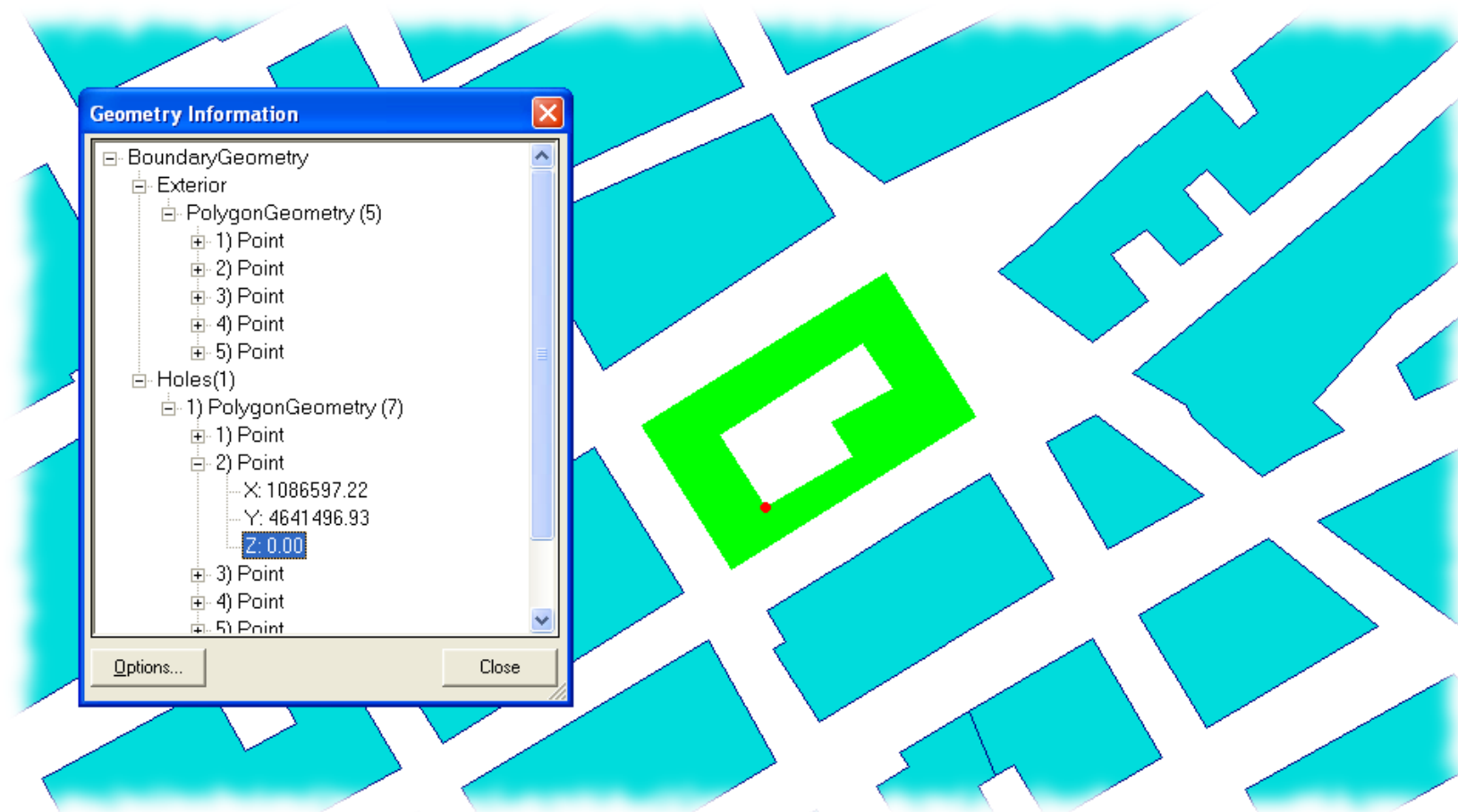
2D e mezzo e 3D

- Cosa si intende per 2D e mezzo?
- La possibilità di memorizzare la Z (altimetria), senza però che questa entri nei meccanismi di analisi ed interrogazione.
- Es: posso fare facilmente ricerche per intervalli di latitudine e longitudine, ma non intervalli per altimetria.
- Ragione di ciò: le analisi geografiche sono fondamentalmente bidimensionali.
- Es: la topologia dei fiumi è spesso corretta sul piano e scorretta in altimetria (gli attacchi in quota degli affluenti non sono coerenti).

Full 3D: Google Earth 3D Buildings



Geomedia: Geometria areale



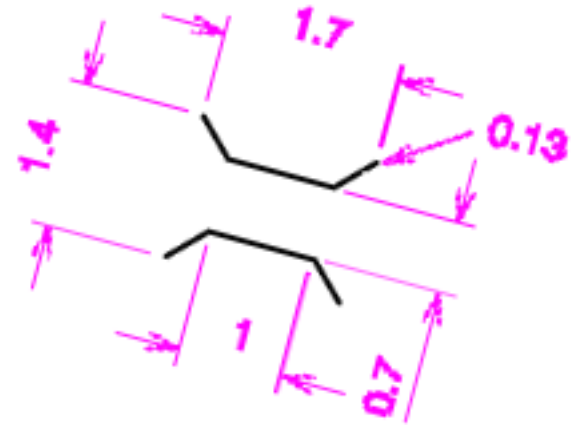
Punto (Point)

- Singolo oggetto puntuale
- E' definita solo la posizione
- Dati: x,y,z
- Es. punto quota



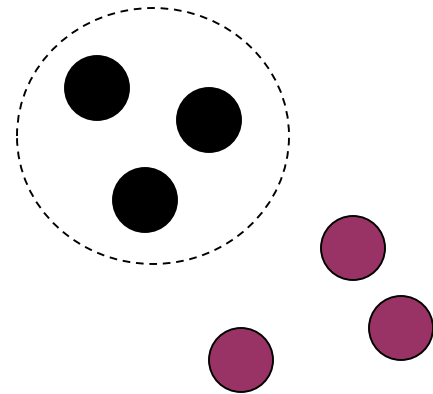
Punto Orientato (Oriented Point)

- Punto con indicazione dell'orientamento
- Dati: $x, y(,z)$ +
 - Angolo piano
 - Versore I, J, K
 - Quaternione
- Es. ponte o aeroporto in scala 250k (è molto importante l'orientamento della pista).



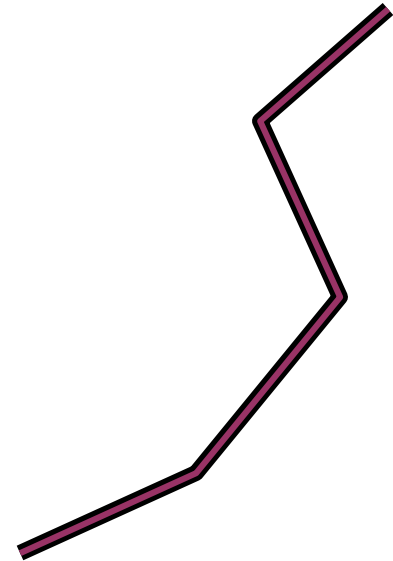
Multi Punto (MultiPoint)

- Oggetto unico definito da una collezione di punti separati.
- Dati, x_1, y_1, z_1 , x_2, y_2, z_2
- Es. Bosco, agglomerato urbano



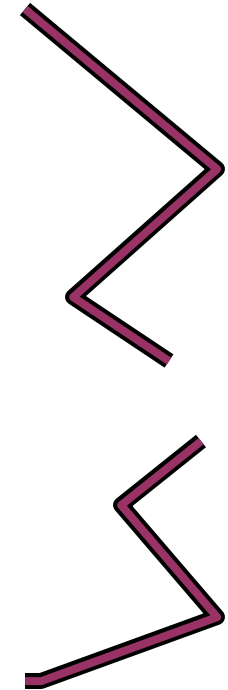
Linea Spezzata (Polyline)

- Curva definita da una serie connessa di segmenti.
- Dati, $x_1, y_1(,z_1)$, $x_2, y_2(,z_2)$
- Es. strada, fiume



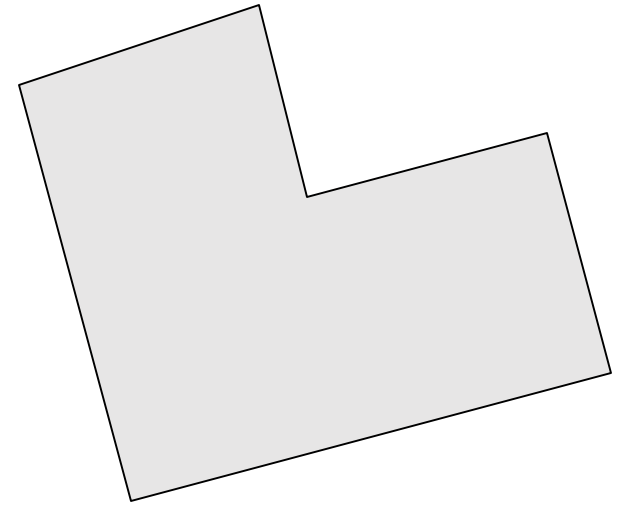
Multi Linea (MultiPolyline)

- Insieme di più curve disconnesse.
- Dati, $x_{11}, y_{11}(,z_{11})$, $x_{12}, y_{12}(,z_{12})$
- $x_{21}, y_{21}(,z_{21})$, $x_{22}, y_{22}(,z_{22})$
- ...
- Es. ferrovia interrotta



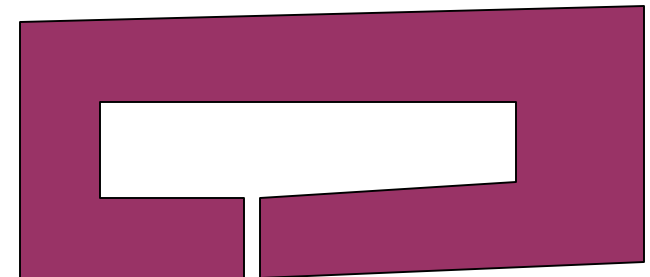
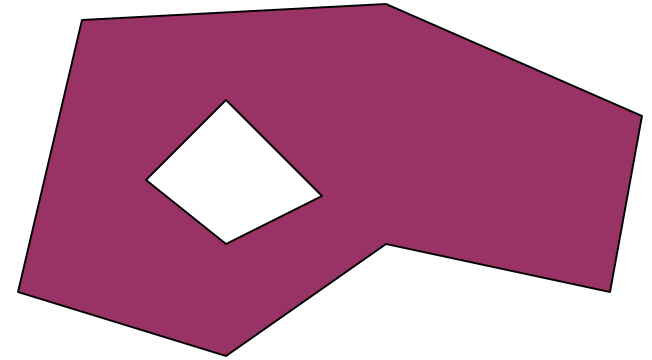
Area Semplice

- Singolo Poligono connesso (concavo o convesso).
- Dati: $x_1, y_1(, z_1)$, $x_2, y_2(, z_2)$
 - Dubbio: ultimo vertice ripetuto?
- Es. Edificio



Area Con Buchi (Area with Holes)

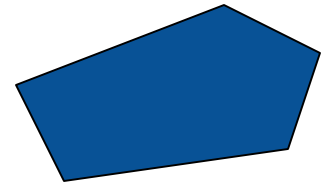
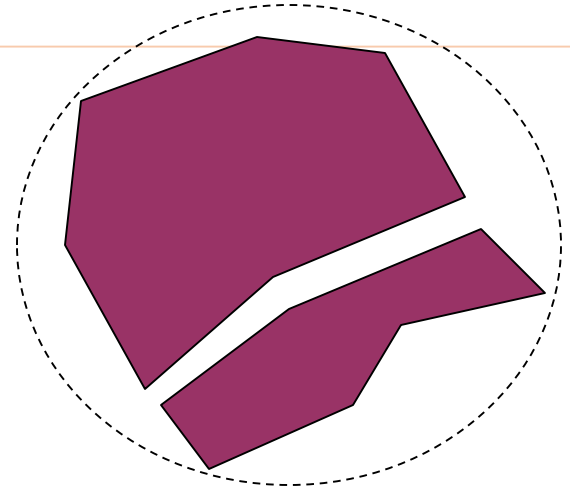
- Singolo Poligono che può contenere uno o più buchi di forma poligonale.
- Dati, bordo (poligono) + lista dei buchi (poligoni)
- Es. casa + cortile, bosco con radure, lago con isole, etc.



Cortile in Microstation

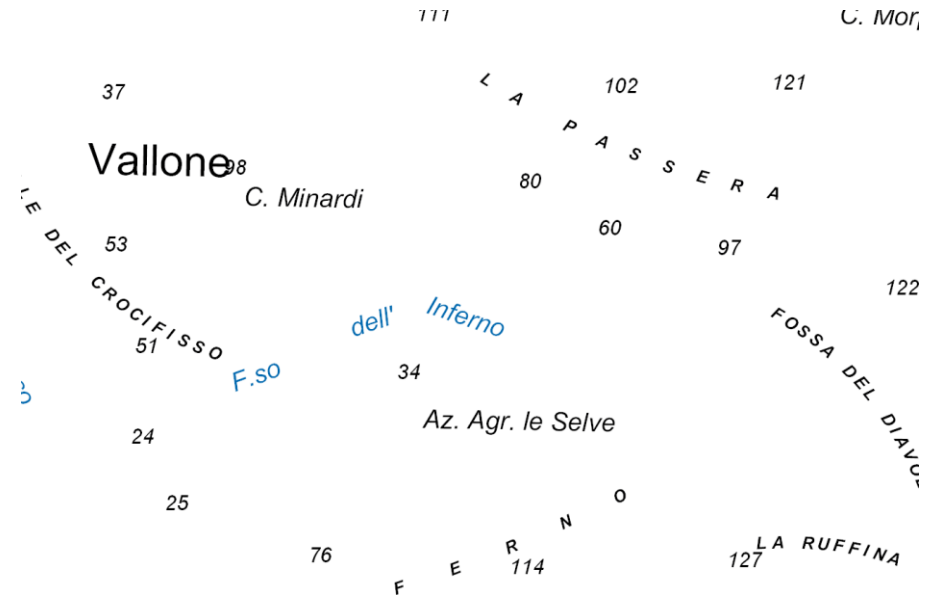
Multi Area

- Oggetto formato da un'insieme di aree sconnesse fra di loro (con buchi o no).
- Dati: lista di aree
- Es. bosco tagliato da un'autostrada.



Testo (Cosa è?)

- Il testo è un oggetto geometrico con del testo attaccato.
- Ma cos'è questo oggetto geometrico?



Problemi

- Aree Microstation senza buchi: importazione da Geomedia?
- Testi: da Punti Geomedia a Polilinee VPF, lunghezza della polilinea??
- Regione Toscana (File ArcView): orientamento dei punti (ponti, etc.) perduto.
- Orientamento DB25 in VPF, Shape Oracle, soluzione: attributo db angolo (AO1)
- DB regionali complessi (join-multigeometrie)

Caratteristiche dei Dati Vettoriali



Le coordinate

- A parte il sistema di riferimento, le coordinate possono essere rappresentate in vari modi:
 - Numeri interi (es. Microstation, Personal GeoDB di ArcGIS fino alla versione 9.0)
 - Numeri in virgola mobile a singola o doppia precisione (es. Shp, Personal GeoDB di Geomedia o di ArcGIS da 9.1 in poi)
- L'importazione/esportazione fra i dei sistemi può introdurre degli effetti collaterali (perdita di precisione, perdita della topologia, spostamento dei vertici)
- La precisione non dipende dal formato ma dalla fonte di acquisizione o dalle specifiche di qualità

Cenni di Topologia

- Possibili livelli topologici:
 - Spaghetti: nessuna topologia
 - Grafi: quale strada è connessa a quale incrocio:
programmi di ricerca percorso su strada
 - Full: informazione esplicita e completa. Es. a destra della
strada x è presenta la coltura y
- Negli ultimi tempi le strutture topologiche complete hanno perso importanza. Ma utili in applicazioni particolari: es. derivazione di scala

Attributi

- Nei dati CAD la sola forma geometrica caratterizza il dato (eventualmente anche le impostazioni di visualizzazione)
- Nei GIS gli oggetti sono caratterizzati anche da attributi associati
- Gli attributi derivano dalle basi di dati: possono essere testi, numeri, date, valori vero/falso, etc.
- Ad esempio una strada può avere i seguenti attributi: nome, larghezza, tipo di asfaltatura, livello al suolo (sopraelevata, sottopasso, etc.)
- Gli attributi sono importanti come la geometria (es. le specifiche di qualità ne prevedono i livelli di correttezza)

Gerarchia in Classi

- Gli oggetti vettoriali sono raggruppati in classi: Layers o Feature Class di oggetti omologhi. Es. strade, fiumi, edifici, etc.
- In lingua italiana possiamo dire “strati informativi”
- In alcuni sistemi gli oggetti di una classe devono essere tutti dello stesso tipo geometrico (es. shape, Personal GeoDB), in altri questo vincolo non esiste (Oracle,Postgres)
- In tutti i sistemi gli oggetti di una classe hanno insiemi di attributi identici (non i valori, ma i tipi ed i nomi degli attributi)
- Nelle basi di dati le classi corrispondono alle tabelle del db

Altri livelli Gerarchici

- La divisione in Feature Class o Layers è obbligatoria
- Alcuni sistemi possono prevedere o no altri livelli gerarchici:
 - Shape e Geomedia: niente
 - ArcGIS: feature-dataset
 - VPF: biblioteche-layers-feature class-tiles
- I livelli gerarchici possono essere impliciti nei nomi delle feature (Specifiche Intesa-Regioni) oppure impliciti solo logicamente (Tematismi DB25)

Gerarchia National Core

**Specifiche Intesa Stato Regioni - Istituto Geografico Militare**

00-Informazioni Geodetiche, Fotogrammetriche E Di Metainformazione
[01-Informazioni Geodetiche](#)
[02-Informazioni Cartografiche](#)
[03-Informazioni Fotogrammetriche](#)

01-Viabilità, Mobilità E Trasporti
[01-Strade](#) 
[02-Ferrovie](#)
[03-Altro Trasporto](#)

02-Immobili Ed Antropizzazioni
[01-Edificato](#)
[02-Manufatti](#)
[03-Opere Delle Infrastrutture Di Trasporto](#)

[04-Opere Di Sostegno E Di Difesa Del Suolo](#)
[05-Opere Idrauliche, Di Difesa E Di Regimazione Idraulica](#)

03-Gestione Viabilità E Indirizzi
[01-Toponimi E Numeri Civici](#)
[02-Amministrazione Viabilità](#)

04-Idrografia
[01-Asse Intesa E Di](#)

Strato 01: Viabilità, Mobilità E Trasporti

Tema 01: Strade

Le classi del Tema

- Classe 01: AC_VEI - [Area Di Circolazione Veicolare](#)
- Classe 02: AC_PED - [Area Di Circolazione Pedonale](#)
- Classe 03: AC_CIC - [Area Di Circolazione Ciclabile](#)
- Classe 04: AR_STR - [Area Stradale](#)
- Classe 05: AR_VMS - [Viabilità Mista Secondaria](#)
- Classe 07: EL_STR - [Elemento Stradale](#)
- Classe 08: GZ_STR - [Giunzione Stradale](#)
- Classe 09: TR_STR - [Tratto Stradale](#)
- Classe 10: IZ_STR - [Intersezione Stradale](#)
- Classe 12: EL_CIC - [Elemento Ciclabile](#)
- Classe 13: GZ_CIC - [Giunzione Ciclabile](#)
- Classe 14: RT_ST1 - [Rete Stradale Liv.1](#)
- Classe 15: RT_ST2 - [Rete Stradale Liv.2](#)
- Classe 16: EL_VMS - [Elemento Viabilità Mista Secondaria](#)
- Classe 17: GZ_VMS - [Giunzione Di Viabilità Mista Secondaria](#)
- Classe 18: RT_VMS - [Rete Della Viabilità Mista Secondaria](#)
- Classe 19: RT_CIC - [Rete Ciclabile](#)

Il sistema di riferimento

- Numeri e coordinate nei GIS non bastano: bisogna indicare un sistema di riferimento
- In cosa consiste: di solito è un codice (numero intero) che riferisce una tabella di Sistemi più o meno standard (es. EPSG).
- Ci possono essere incongruenze o incomprensioni (es. Oracle ha i suoi codici, Postgres non legge i prj)

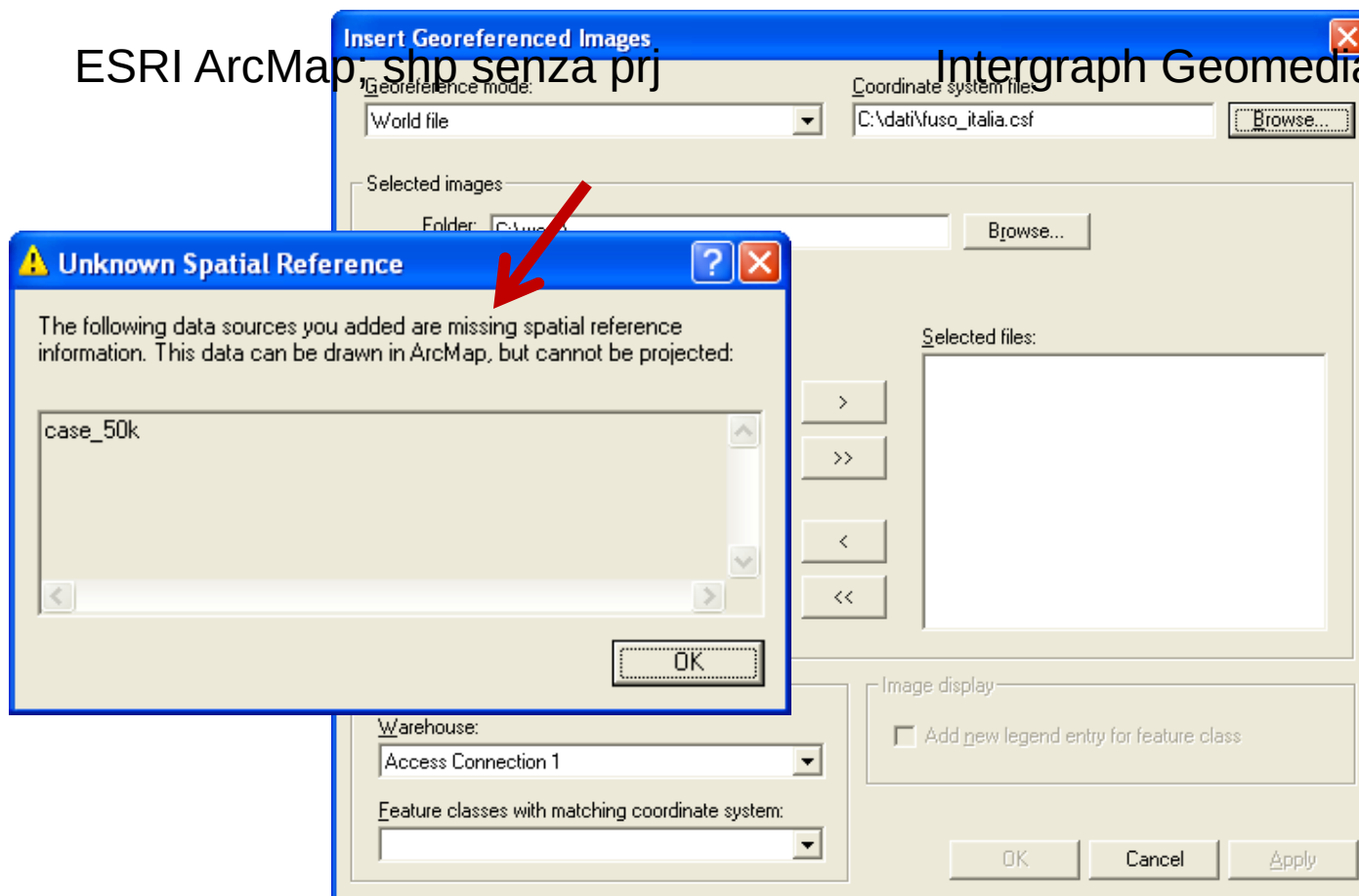
Sistemi di Riferimento

- Vari livelli di informazione
 - Assenza di informazione (tif+tfw, file DXF)
 - Informazione associata esterna (shp + prj)
 - Informazione intrinseca:
 - Per intero database (Geomedia 5)
 - Per feature (shp, Geomedia 6, ESRI Personal DB)
 - Per singolo oggetto (Oracle Spatial, Postgres + Postgis)

Mancanza di Informazione

ESRI ArcMap: shp senza prj

Intergraph Geomedia: TIF+TFW



Formati e Metodi di Memorizzazione



Accenno ai Formati di Dati Vettoriali

- Formati CAD: DXF, DWG, DGN: non contengono attributi (o metodo per associare attributi esterni). Non contengono propriamente Classi ma livelli grafici
- Formato SHP: antico e solido, di origine proprietaria (ESRI), di fatto uno standard.
- SVG: formato xml orientato alla visualizzazione grafica (no Classi ed attributi)
- GML: formato xml orientato al GIS. Tentativo di renderlo uno standard accettato (es. Progetto INSPIRE)
- E00: esportazione file ArcGIS
- Formati Database file-based: Personal GeoDatabase di Geomedia o di ArcGIS

Caratteristiche di Formati

- I vari tipi di formato si distinguono per alcune caratteristiche:
 - Dimensioni supportate (2, 3 o 4).
 - Tipi geometrici supportati: punti, punti orientati, polilinee, aree, aree con buchi, oggetti multipli.
 - Gestione toponimi: label o features, punti, archi di cerchio, polilinee.
 - Presenza di dati associati (database cartografici), o presenza di etichettatura minima.
 - Sistema di riferimento esplicitamente associato: sì o no.
 - Possibilità di rappresentare classi di oggetti (gerarchia piatta).
 - Possibilità di rappresentare gerarchie di classi.
 - Specifiche di formato: protette (Oracle), semiprotette (DGN), pubblico dominio ma proprietarie (Shape), non proprietarie (GML).
 - Altro: es possibilità di inserire metadati espliciti.

Nota sui formati protetti

- La legislazione americana permette di brevettare i formati di file (quella europea è più libertaria).
- Es. formati pubblici come il **tiff compresso** (ma non quello normale) e il **gif** sono brevettati.
- Per il loro utilizzo (anche la sola lettura) è necessario pagare i diritti del brevetto alla società proprietaria.
- Alcuni produttori di software rinunciano a gestire questi formati per non far lievitare i costi dei prodotti.
- Altri formati non hanno specifiche pubbliche e sono secretati. Esistono specifiche non pubbliche (reverse-engineering) che comportano poca solidità o addirittura conseguenze penali.

Classi di formati

- Tabelle dati di coordinate (punti)
 - File testuali, tabelle di fogli elettronici (Excel)
 - Tabelle dati di basi di dati classiche (Access)
- Formati CAD
- Formati Geografici File-Based
- Basi di Dati Cartografiche
 - Personal (DB File-Based)
 - Enterprise (DB Server)

Microstation DWG

- Sistema di riferimento non presente (vecchie versioni), supportato nelle nuove (con tag aggiuntivo).
- Tipi geometrici: punti orientati (simboli), polilinee, aree senza buchi, testi come punti orientati.
- Suddivisione in livelli. Possibilità di collegamento ad un database esterno.
- Limite di 101 vertici alle polilinee ed alle aree (nota da informatico: vincolo veramente limitante ma permette una gestione efficientissima della paginazione su disco).
- Dimensioni: 2 o 3 (comune a tutto il file).
- Formato semi-pubblico (possibile leggere il binario, impossibile crearlo da zero, campi tag segretati).

DXF

- Formato derivato dal CAD (di base geometrie + attributi grafici + layers).
- Ma formato testo molto semplice con specifiche pubbliche: ottimo per la divulgazioni di dati CAD.
- Struttura generale del file testo: codice del valore (tag) seguito dal valore, es: 10 x 20 y 30 z
- Problematiche: grande quantità di estensioni ai tag di base, non del tutto standardizzati (es. coni e cilindri tridimensionali ruotati nello spazio), che rendono difficoltosa la gestione del file.

Shape File

- Formato file based (o meglio multiple file-based: una feature da 3 a 5 file). Gli attributi sono memorizzati in un file dbase III (leggibile anche con Access).
- Formato proprietario di pubblico dominio.
- Tipi geometrici: punti, polilinee, aree, oggetti multipli.
- Dimensioni supportate: 2,3 e 4 (dato misura associato ai nodi).
- Sistema di riferimento non incluso (di default), nelle ultime versioni del software viene associato al file prj (Non supportato da altri produttori: es. Intergraph)
- Semplice e molto famoso.

Geomedia 6 Intergraph

- Basato su database Access, con codifica binaria delle geometrie.
- Formato di pubblico dominio (quasi: Spatial Key secretata).
- Tipi geometrici: punti orientati, polilinee ed aree con buchi, compound esclusivamente 3D, testi come features.
- Classi di feature, nessuna gerarchia.
- Metadati minini (nome alle features).
- Sistema di riferimento per database (V4), per feature (V5).
- Domini codificati di attributi (V5-V6).

ESRI – Personal Geodatabase

- Basato su database file-based Access.
- Tipi geometrici: punti, polilinee, aree con buchi, oggetti multipli, sempre 3D.
- Classi di feature, gerarchia in feature dataset
- Domini per intervallo o codificati.
- Sistema di riferimento per feature (per attributo geometrico), o per feature-dataset. Concetto di estensione e precisione nei metadati.
- Coordinate intere (9.0) o virgola mobile (9.1)

ESRI Export FILE (e00)

- Formato di esportazione dei coverage di ArcInfo (che quindi ne conserva le caratteristiche).
- I coverage hanno topologia esplicita
- Le specifiche non sono ufficialmente pubbliche (non c'è un documento ESRI), ma il formato è stato “craccato” a vari livelli (reverse engineering).
- Supportato in lettura da molti sistemi di visualizzazione.
- Non è possibile definirlo come un vero sistema di divulgazione dei dati.

Mapinfo MIF File

- File di esportazione di Mapinfo
- Semplice file testo con specifiche pubbliche e libere.
- Molto facile da leggere e scrivere.
- Totalmente orientato alla visualizzazione della cartografia, contiene geometrie e simbologia di visualizzazione.
- Non si può parlare di database cartografico.

VPF

- Sistema di database-file based (multiple file based, la struttura è realizzata con le cartelle del file system).
- Geometria: punti, linee, aree, testi. Dimensioni: 2 o 3D.
- Gerarchia di contenuto: DB – Tematismi – Feature Class.
Gerarchia Spaziale: DB – librerie – tiles.
- TOPOLOGIA ESPLICITA (spigolo alato), di vario livello: spaghetti, grafo, grafo planare, full topological.
- Formato pubblico non proprietario.
- La geometria è sempre dissociata dalle features: oggetti diversi possono condividere la stessa geometria (memorizzata una volta sola).

GML

- GML è un'estensione dell'XML (che deriva da HTML) specificata in UML.
- Formato testo di pubblico dominio.
- Permette di rappresentare tutte le informazioni di db: geometrie, attributi, metadati, vincoli.
- Non essendo di origine commerciale, necessita di un forte supporto.
- Base delle specifiche dei servizi web (WFS), e di tutti i formati moderni (es. metadati CNIPA)
- Esiste in più versioni leggermente discordanti

KML (Keyhole Markup Language)

- Nasce da Google Earth (formato commerciale)
- Molto semplice ma molto diffuso
- Permette di rappresentare semplici strati geometrici
- Permette di memorizzare semplicifi informazioni ausiliarie ed etichette.
- Adesso standard internazionale

Notorietà del formato

- Aspetto importante del formato: la notorietà.
- Non è detto che venga usato il formato migliore tecnicamente.
- Ci sono motivi economici, storici, di simpatia personale etc.
- Non si può imporre un formato per direttiva, se il mercato non lo accetta.

Supporto del formato all'efficienza.

- Presente nei formati di lavoro e di utilizzo diretto nella divulgazione.
- Dati secondari derivati (quindi duplicati) dei dati principali, con funzione unica di velocizzazione dell'utilizzo.
- CAD: a volte mbr (minimum bounding rectangle o box).
- Shape: idx file, per il recupero diretto degli attributi.
- Geomedia 4,5: attributo Spatial Key per ordinamento frattale (codifica segreta) + indici testuali Access.
- ESRI Personal GeoDB: Tabella Shape_Index (mbr) + indici numerici Access.
- VPF: indici spaziali efficienti espliciti (Quad-tree), indici espliciti sugli attributi, indici Hash per la ricerca veloce di sottostringhe.
- MySql: indici spaziali grossolani sul mbr.
- PostgreSQL: indici spaziali efficienti, da selezionare manualmente (RTree).
- Oracle: indici spaziali altamente efficienti, selezionati automaticamente dal sistema (tecnologia protetta: probabilmente z-ordering frattale).

Divulgazione tramite Web

- In questo caso non si parla di un formato ma di un software che permette di creare siti web dinamici basati su cartografia vettoriale.
- Il formato utilizzato rimane nascosto all'utente (es. oracle spatial o shape).
- Software ESRI: ArcIMS
- Software OpenSource: Mapserver (NASA)
- Il futuro è questo: servizi WEB (WFS WMS)

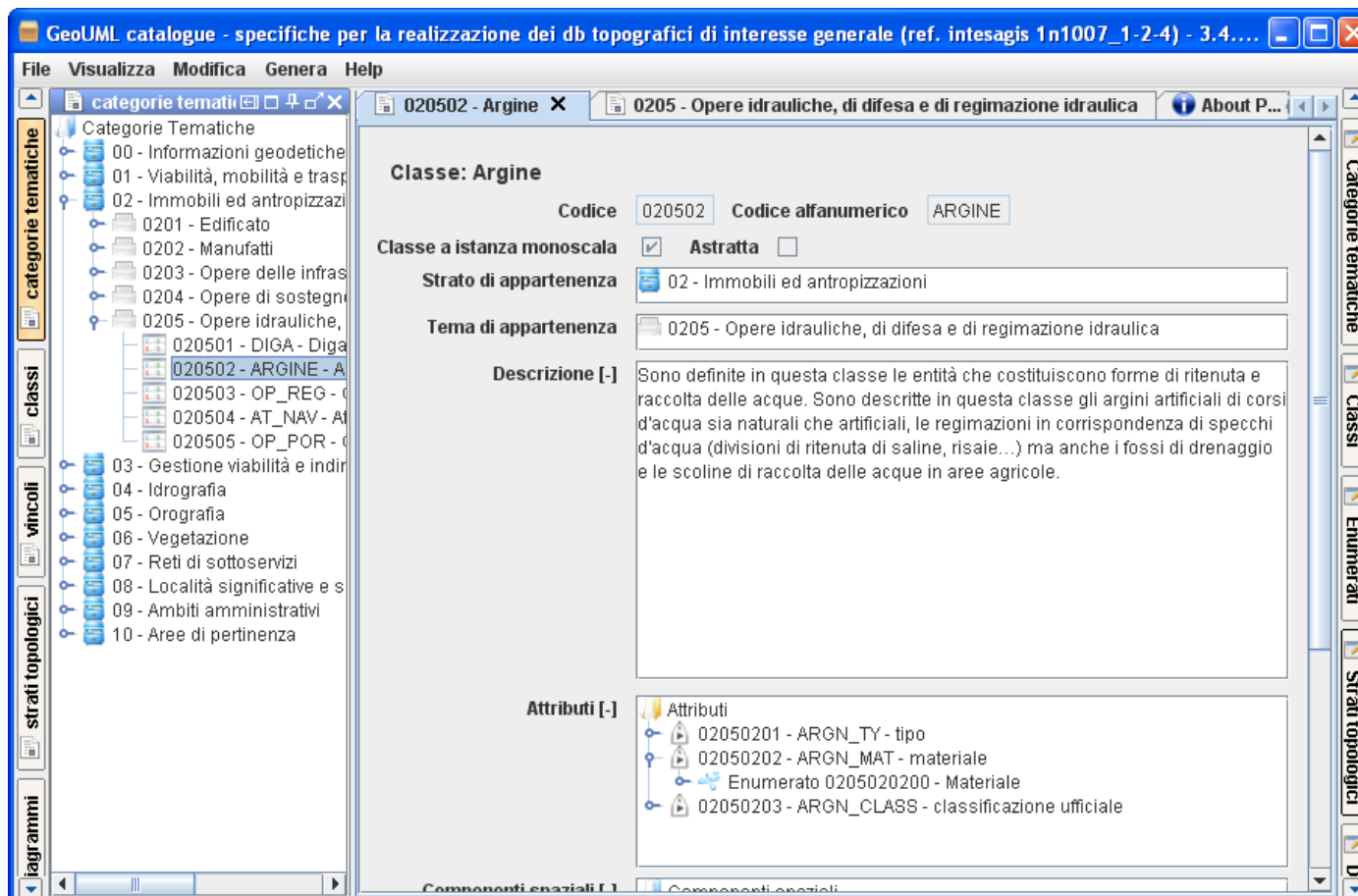
Problema di Standard

- I formati vettoriali hanno avuto una storia più sfortunata di quelli raster
- Non esiste uno standard libero accettato di fatto per lo scambio dati (GML è un tentativo)
- Molti formati sono proprietari (di ditte private), es. shp, oppure sono addirittura semi-secretati (DGN Microstation) o totalmente secretati o coperti da brevetto
- Questa situazione comporta molte problematiche: es. capitolati di appalto, scambi di dati ufficiali.
- Il problema potrebbe essere risolto dall'abbandono di formati file ed il passaggio a servizi di rete WMS, WFS, che non prevedono la presenza di formati di file.

Specifiche Ufficiali

- Il dato vettoriale è molto complesso: ha bisogno di specifiche molto articolare
- Le specifiche possono essere a più livelli (modello logico, modello fisico)
- Di solito descrivono:
 - le classi di oggetti: da cosa sono composte, che tipo di geometrie le formano
 - l'elenco degli attributi: i tipi di dato, i limiti dei valori oppure i domini espliciti (elenco valori possibili)
 - Le relazioni ed i vincoli topologici
- Le specifiche si presentano spesso sotto forma di documenti (pdf): sarebbe meglio che esse stesse siano strutturate in basi di dati per poter effettuare collaudi automatici
- Es: XML schema dei metadati CNIPA: permette il controllo di correttezza automatico di una scheda metadati.
- Es. IGM collaudo puglia: cattura automatizzata delle informazioni dai documenti ufficiali di specifica -> generazione database delle specifiche (metedatabase) -> controllo automatico dei dati (struttura, domini, vincoli topologici)
- Il discorso sarebbe lungo...

Specifiche Politecnico di Milano



NATO DGIWG FAD

Resource Link [Short Name]	5-3-1 Code	Name	Definition	Description
AccessZone	FA005	Access Zone	A terrain region between a contact zone and the first passable land transportation route (for example: a road).	
AcousticStation	BK010	Acoustic Station	The geographic location at which a set of acoustic observations were taken.	
ActorEntity	ZI044	Actor Entity	An abstract modeling entity that is a superclass for actor types, which are representations of intentional (that is, purposeful) entities that act or have the capability of acting as a participant in an event.	An actor may control resources or direct activity, and may do so by holding an organizational role. Actor entities include persons, social groups, organisations, and, in some contexts, automata. Collective actors (for example: organizations) may have a distributed rather than a point location.
AdministrativeBoundary	FA000	Administrative Boundary	A boundary between administratively controlled regions.	
AdministrativeDivision	FA003	Administrative Division	An administratively subordinate division of a geopolitical entity.	A geopolitical entity (country) is typically divided into first-, second-, and lower-order administrative divisions. First-order administrative divisions are immediately subordinate to the government of the geopolitical entity, with second- and lower-order divisions subordinate to those above them. Examples: (first-order) a United States state, a German Land, a French region; a Canadian province; (second-order) a U.S. county, a French department; (third-order) a U.S. township, a French arrondissement; (lower-levels) a French commune.
AdminDivisionDesig	ZI021	Administrative Division Designation	A modeling entity collecting information about designations (for example: standardized names or codes) for a subordinate administrative division (for example: a (US) state, a (UK) county, a (US) township, a (CA) province, a (FR) arrondissement, a (CH) canton, a (GE) laender and a (FR) commune) of a geopolitical entity (for example: a State).	
AdministrativeRegion	FA001	Administrative Region	A region controlled by an administrative authority.	
AerationBasin	AB040	Aeration Basin	A basin, usually artificial, in which air is mixed with partially treated wastewater.	The water is treated with microorganisms to consume organic materials and convert suspended solids to settleable solids that are later collected by a settling pond and removed as sludge.
Aerial	AT011	Aerial	A device that is used for emitting and/or sensing electromagnetic energy.	For example, used to transmit and/or receive electronic signals as on a radio tower or to capture electromagnetic energy as in radio astronomy.
AerialCablePylon	AQ020	Aerial Cable Pylon	A pylon supporting an aerial cable.	For example, a ski lift pylon.
AerialFarm	AT012	Aerial Farm	A collection of aerials that are collocated and serve a common purpose.	They may be organized either to function as a single larger virtual device (for example: a phased array) or function relatively independently of each other (for example: pointed in different

I database Spaziali



Database Server

- Nei database professionali (Oracle, Postgres) i dati non corrispondono a file
- In questo caso non si può parlare di un formato vero e proprio
- Per “trasportare” i dati di questi database non si può copiare i file direttamente, ma:
 - Si possono utilizzare meccanismi di esportazione / importazione che trasformano i dati in file DUMP
 - Si possono utilizzare tool specifici per la comunicazione con altri formati, es. shp2pgsql o shptosdo.

DB – Server: caratteristiche

- Centralità ed unicità del dato (in contrasto con i file-based).
- Accesso condiviso ai dati, anche in fase di modifica: gestione della coerenza di accesso in ogni caso.
- Gestione di enormi quantità di dati, come se fossero pochi (in teoria, quasi indipendenza dalla cardinalità dei dati).
- Non corrisponde a file su disco.

MySQL 5

- Possibilità di creare colonne geometriche.
- Tipi geometrici supportati: punti, polilinee, aree, oggetti multipli, ma solo in due dimensioni.
- Nessuna gestione effettiva del sistema di riferimento.
- Limitato insieme di query spaziali, limitatissimo insieme di funzioni di manipolazione: previsto insieme completo ma non ancora implementato. *“The OpenGIS specification defines the following functions, which MySQL does not yet implement. They should appear in future releases”.*
- Inportazione/esportazione: secondo le specifiche OGC, Well kwon format sia binario che testo.
- Non supportato dai sistemi commerciali. Non più supportato dai sistemi OpenSource (perché uscito dalla licenza GPL).
- Molto veloce in query di lettura, molto lento in modifica ed accesso condiviso.

PostgreSQL - PostGIS

- Possibilità di creare attributi spaziali come veri attributi, le feature possono avere quindi:
 - Più di una rappresentazione geometrica per la stessa feature.
 - Geometrie miste all'interno della stessa classe.
- Tipi geometrici:
 - Punti, linee, aree (anche con buchi).
 - Multi oggetti e compound misti
- Sistema di riferimento per feature
- Conversione di sistemi di riferimento.
- Vasta possibilità di query spaziali e di funzioni geometriche.
- Non automazione nell'uso degli indici spaziali: necessita di pratica nella scrittura di query.
- Esportazione/Importazione: testuale e binaria secondo le specifiche OGC! (Well know format). File shape
- Mal supportato dai software commerciali per motivi di cartello (ma utilizzabile da Mapserver e GRASS, e ArcGIS 9.1 con Interoperability extensions).
- Media velocità in lettura, molto veloce in modifica ed accesso condiviso.

Oracle Spatial 11g

- Possibilità di creare attributi spaziali come veri attributi, le feature possono avere quindi:
 - Più di una rappresentazione geometrica per la stessa feature.
 - Geometrie miste all'interno della stessa classe.
- Tipi geometrici:
 - Punti, linee, aree (anche con buchi).
 - Multi oggetti e compound misti
- Possibilità di specificare il sistema di riferimento sia per classe, sia per singolo oggetto (sistemi diversi all'interno della stessa classe).
- Possibilità di memorizzare metadati sui campi geometrici (es. risoluzione piana).
- Vasta gamma di query spaziali applicabili ai dati.
- Gamma di funzioni di trasformazione (es. baricentro, intersezioni, etc.)
- Supporto completo di indicizzazione spaziale, applicazione automatica dell'indice ottimale per le query spaziali.
- Possibilità di trasformazioni fra sistemi.
- Importazione/esportazioni di dati geografici in formato testo proprietario.
- Completamente Supportato da Intergraph ed ESRI.

Confronto Oracle - Postgres

- Miei pareri personali, ma derivati dall'esperienza (IGM ha 3 server Oracle e 5 server Postgres), il confronto è arduo, la discriminante è il costo.

	Oracle	Postgres
Costo	200000 – 300000 euro	0 euro
GIS Open Source	Mal supportato	Pienamente supportato
GIS Commerciali	Pienamente Supportato	Mal o Non supportato
Supporto (manuali, consulenza)	Perfetto	Scarso od a pagamento
Efficienza	Ottima per dati giganti	Buona – Ottima dati medi
Solidità	Ottima	Scarsa (crash da dati giganti)
Gestione	Ostica (necessita personale specializzato)	Facilitata (personale autodidatta)
Compatibilità Standard	Bassa (deriva da una lunga storia)	Alta (OCG compliant)
Aggiornabilità	Ottima (cambio versione automatico)	Molto faticosa

Approfondimenti



I dati vettoriali e la SCALA

- Per i raster la scala è rappresentata dalla risoluzione a terra (es. 1 pixel = 5 m)
- Per i vettoriali la definizione di scala (sicuramente presente!) è discussa:
 - Ovviamente un dato vettoriale può essere zoommato e ridotto a volontà ma la scala incide su...
 - La densità degli oggetti presenti
 - La precisione di rilevazione delle coordinate
 - La densità dei punti che compongono le geometrie
 - Il tipo di geometria utilizzato per un oggetto (es. singoli edifici areali che diventano un singolo punto centro abitato)

Vettoriali e SCALA (2)

- Per i DB vettoriali quindi si definisce comunque la scala
- Es. DB cartografia tecnica regionale 1:5000, 1:10000: piscine areali, strade areali
- Es. DB25 IGM carta topografica a 1:25000: piscine puntuali, strade lineari
- Es. un tornante di montagna, sempre lineare, avrà forma diversa nelle varie scale, altrimenti non sarebbe fruibile
- La precisione del dato invece è un termine in discussione (alta precisione anche a scale a grande denominatore?)

Perché db a varie scale?

- Perché non basta una sola scala?
 - Motivi di efficienza (impossibile caricare il 2k di tutta una regione)
 - Di punti di vista: se analizzo l'Italia intera non mi interessano i balconi degli edifici (i riassunti sono informazioni importanti)
- Ma:
 - Possibili db multiscala (es. centri abitati 2K e periferia a 10k)

Aspetti Correlati: Simbolizzazione

- Tutte le volte che voglio vedere un DB, devo “vestire” le geometrie
- Si definisce una certa simbologia (es. colore, spessore, doppio bordo, tratteggio, simboli predefiniti)
- Di solito la simbologia è associata ad una Classe di oggetti, oppure all’interno di una classe varia al variare degli attributi
- Per l’IGM questo è un aspetto veramente importante: la simbolizzazione è eseguita con grande cura e non in modo automatico.

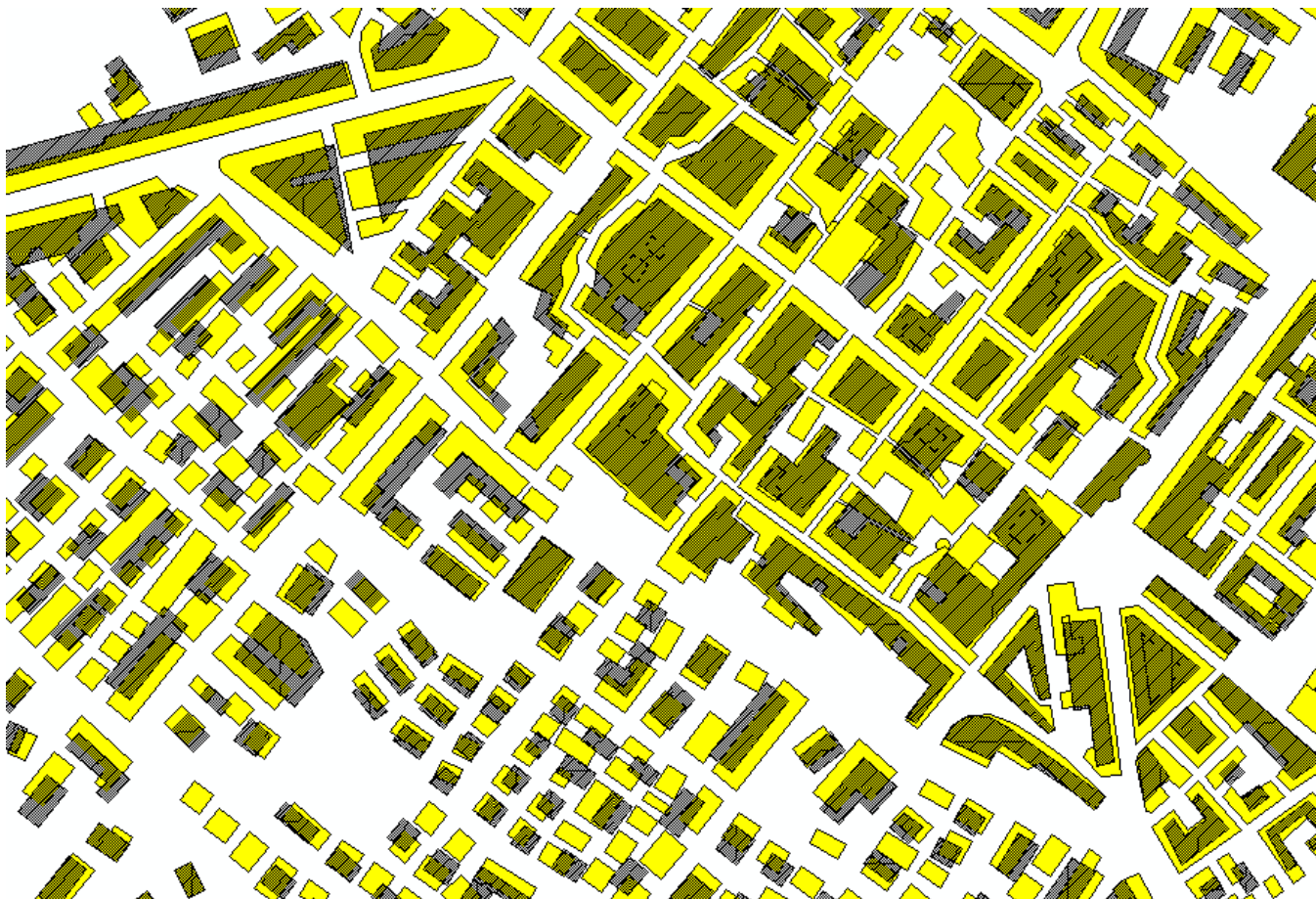
Generalizzazione

- Per generalizzazione si intende il cambio di scala di un dato vettoriale (spesso da piccolo denominatore a grande)
- Aspetto molto importante. Es. le regioni producono carte tecnica a piccolo denominatore: necessità di produrre la cartografia alle altre scale
- Esempi di operazioni:
 - Sfoltimento di oggetti
 - Sfoltimento dei vertici degli oggetti
 - Cambio di tipo geometrico (es da area a punto)
 - Reinterpretazione (da case a centro abitato, da campo da tennis + spogliatoio a centro sportivo, etc).

Generalizzazione 2

- In passato si è tentato di costruire software totalmente automatici, ma il tentativo è fallito
- Spesso è presente un aspetto “artistico” di interpretazione del cartografo che riesce ad riassumere il dato capendolo e reinterprelandolo
- La tendenza attuale è quella di fornire una serie di strumenti software utili (vedere quelli di *ArcGIS* o *Radius Clarity*), che l'operatore può utilizzare interattivamente.
- La parte automatizzata può essere comunque notevole: si ottiene un risparmio del 75% del tempo.

Derivazione 25 -> 50 (IGM)



Derivazione: Nuova Filosofia

- In Arcgis 9 gli strumenti di derivazione erano nel menù “Data Management”
- In Arcgis 10 sono stati spostati in “Cartography”
- Mentre nel passato la derivazione riguardava i dati vettoriali e il db associato, nella nuova filosofia la derivazione sembra riguardare solo la simbolizzazione (stampa).

Object Oriented

- In alcuni casi si parla di dati (o di database) strutturati in modo Object Oriented
- La terminologia ed i concetti (es. incapsulamento, ereditarietà, polimorfismo) derivano dai linguaggi di programmazione (C++, Java)
- Non ci sono implementazioni significative quindi possiamo ignorare questo aspetto
- Però:
 - Alcune specifiche (es. UML) utilizzo però definizioni O.O.
 - Alcuni formati (GML) nascono O.O.

Errori e vincoli

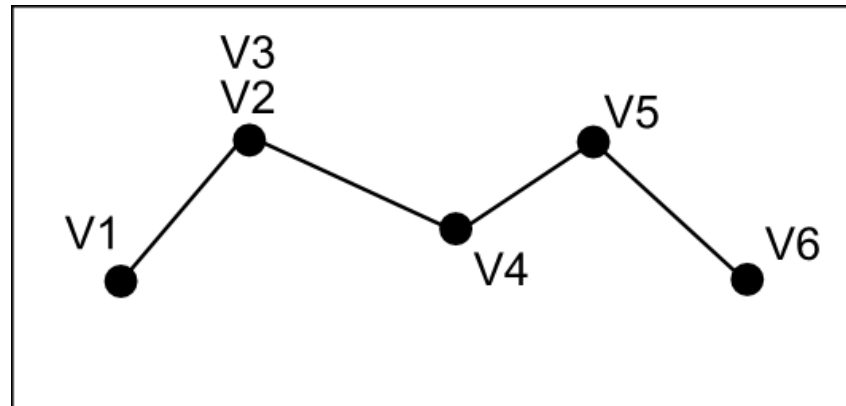


Errori Geometrici

- Le geometrie dei dati vettoriali descrivono oggetti geografici
- Non tutte le rappresentazioni sono corrette: è possibile definire una serie di regole di correttezza per i dati vettoriali
- Fonti di errore possono essere: errori in restituzione, in elaborazione, nel cambiamento del sistema di riferimento
- Molti di questi errori non sarebbero significativi per la stampa. Sono invece deleteri per i metodi di elaborazione ed interrogazione

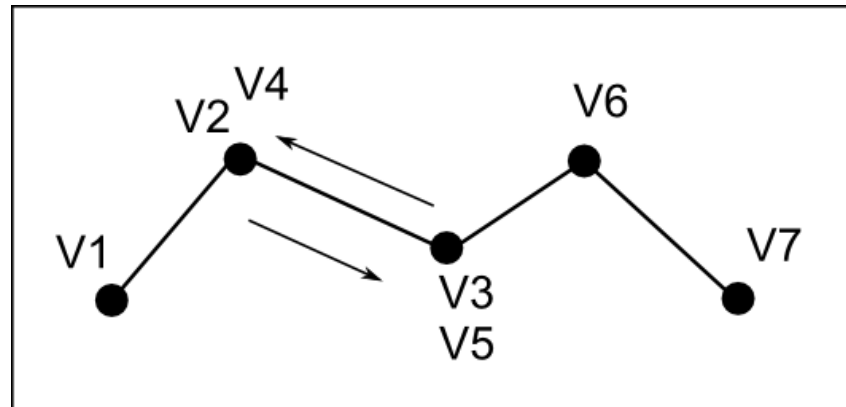
Punto Duplicato

- Tipica Fonte: restituzione
- Comporta errori negli algoritmi di ricerca o nei controlli topologici



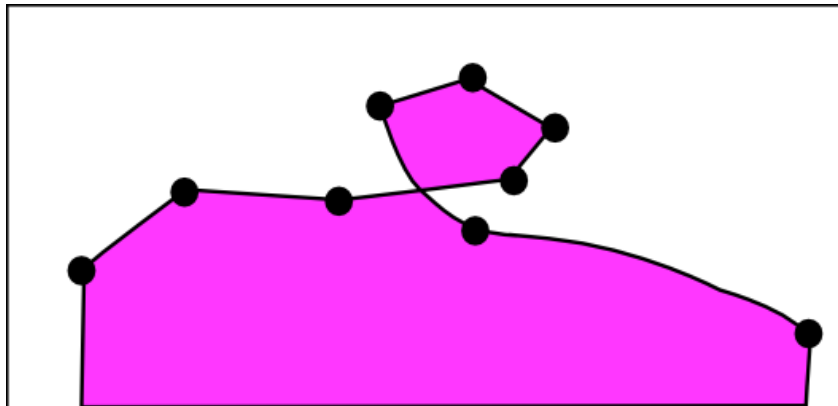
Kick Back

- Fonte Tipica: restituzione
- Non è visibile in alcun modo
- Genera errori come il punto duplicato



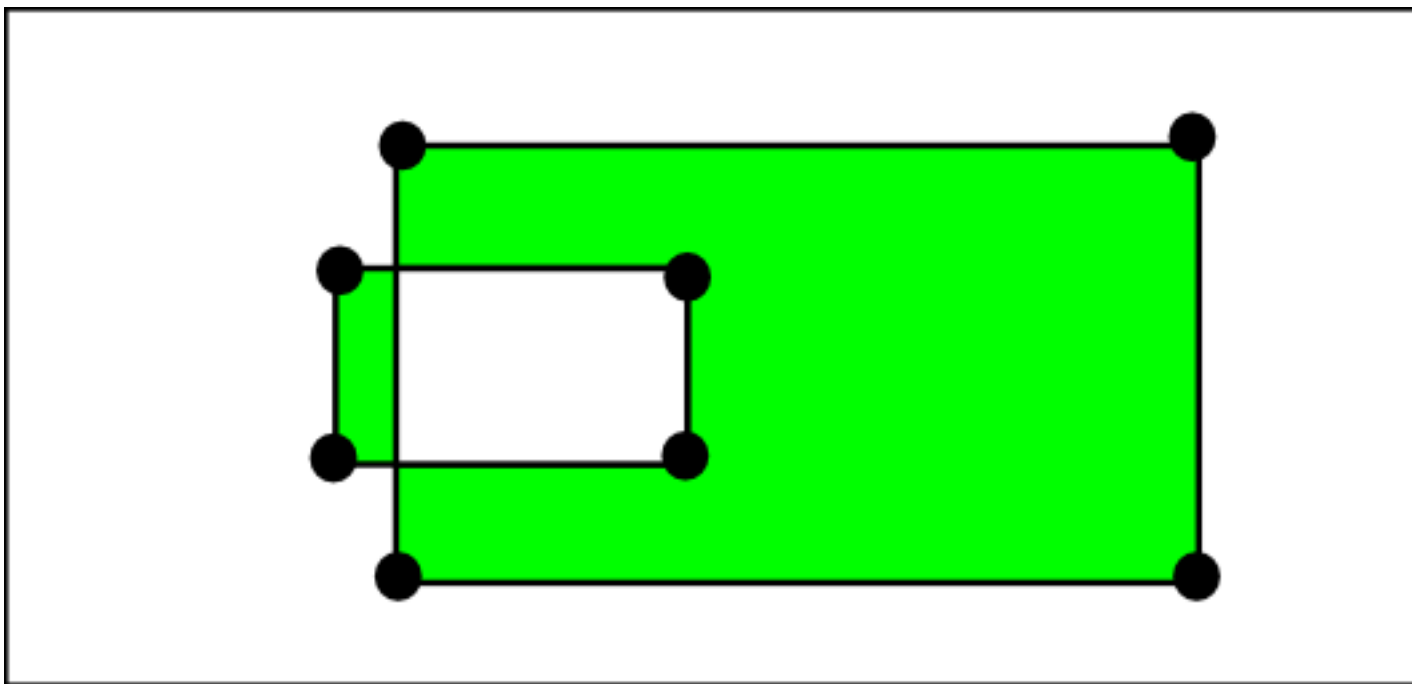
Loop

- Varie Fonti
- Il bordo si auto-interseca: fallimento degli strumenti topologici
- Può essere anche molto piccolo e non visibile



Buchi non inclusi

- Il buco interno non è contenuto nell'area
- Es. fonte: cambio di sistema di riferimento

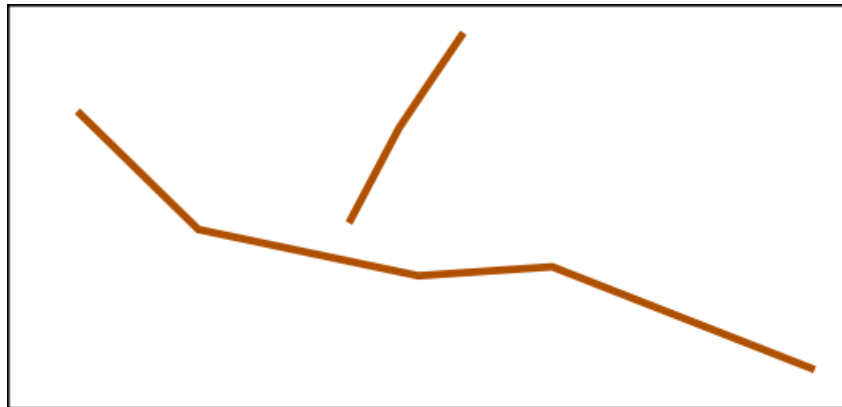


Errori topologici

- Anche non in presenza di topologia esplicita, alcune classi di oggetti sono sottoposte a vincoli topologici di correttezza
- Questi vincoli spesso riguardano una coppia di Classi (es. strade e ponti), oppure coppie di oggetti facenti parte della stessa Classe (es. due strade)
- Il controllo di questi vincoli genera una serie di errori topologici.
- Questo tipo di errori non ha nessun impatto nella visualizzazione o nella stampa: riguarda i metodi di ricerca tipici del dato vettoriale.

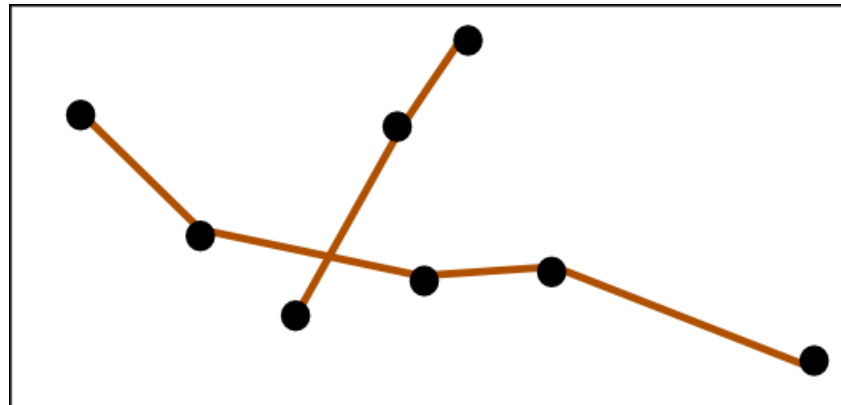
Under shoot

- Es. incrocio a T di strade. Una strada arriva “corta”
- Può essere molto piccolo (millimetri del mondo reale) ma comporta il fallimento degli algoritmi di ricerca



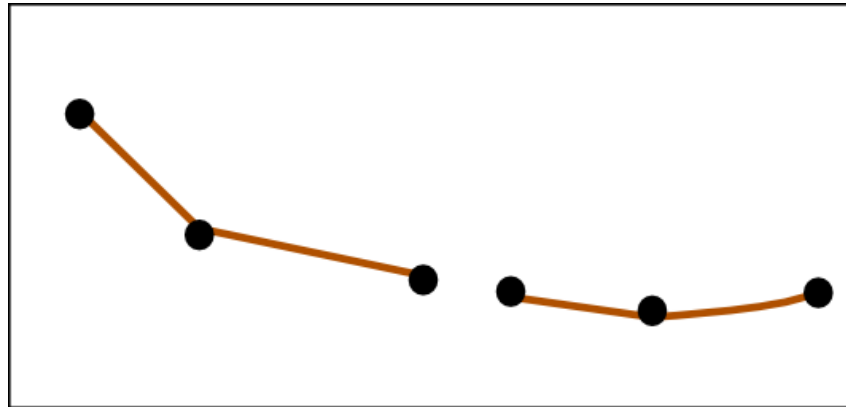
Over shoot

- La strada arriva “lunga”: non c’è un vertice nell’intersezione: le strade risultano scollegate



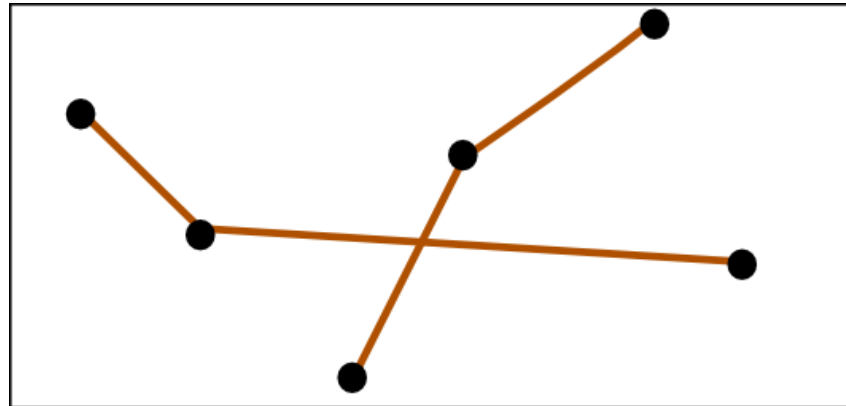
Node Mismatch

- Un piccolo buchino: fallimento algoritmi di navigazione, la strada non è percorribile



Unbrocken geometry

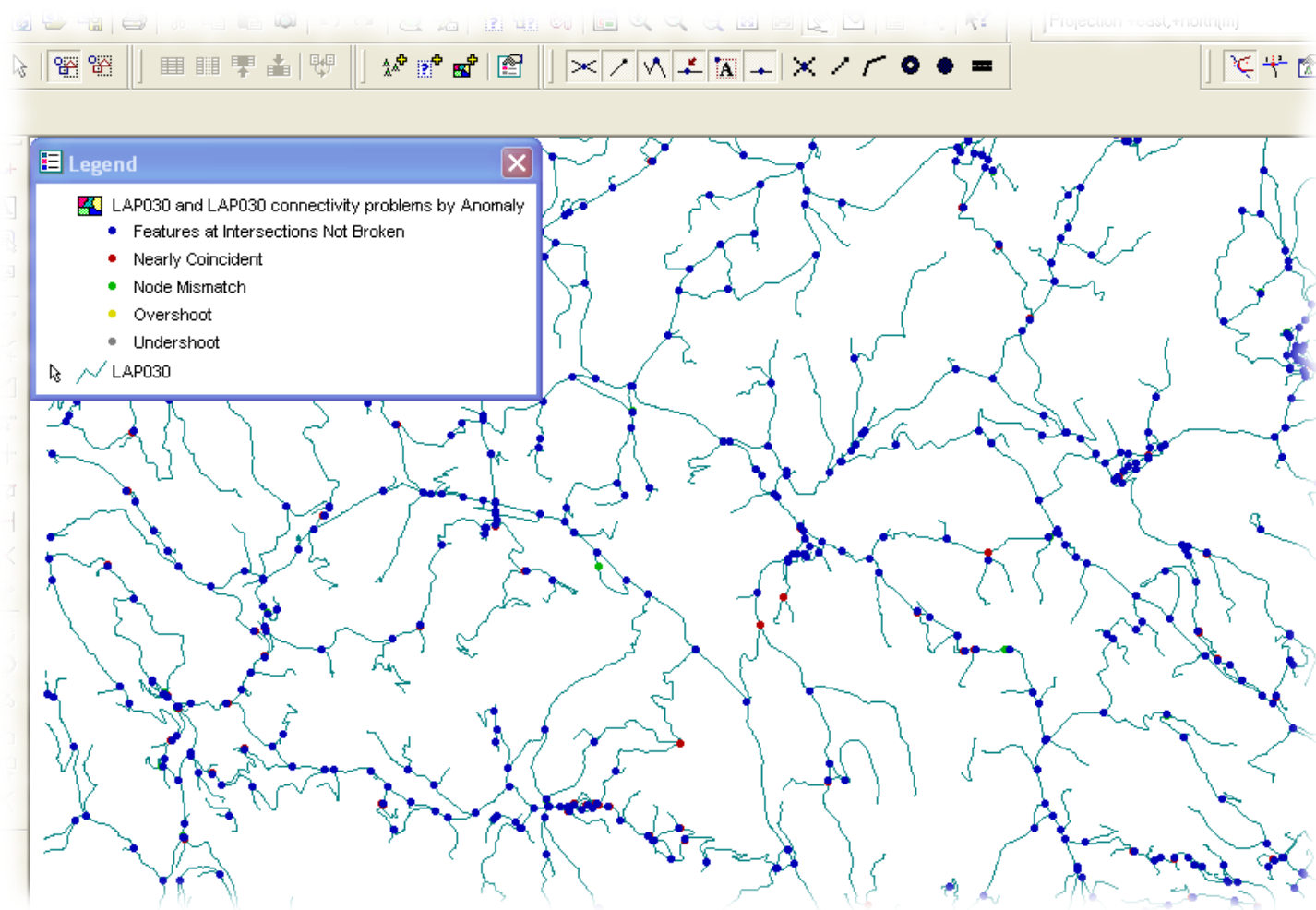
- Un incrocio di quattro strade: manca il vertice comune, quindi le due strade non sono connesse
- Alcune specifiche richiedono che in caso di incroci, gli oggetti siano tagliati (quattro strade distinte)
- Non sempre un errore: corretto in caso di cavalcavia. In questo caso un controllo automatico deve indagare gli attributi degli oggetti.



Altri vincoli Geometrici

- Le specifiche di una particolare serie cartografica possono includere molti vincoli topologici.
- Esempi notevoli:
 - Specifiche cartografia tecnica regione puglia, descritte in modo informale (ma rigoroso in senso matematico) negli allegati del capitolato tecnico
 - Specifiche DB25, in questo caso sono state individuate 45,000 relazioni da rispettare

Geomedia: Analisi Connettività



Utilizzo del dato Vettoriale



Geoprocessing: Processo di Produzione

- Acquisizione: foto-restituzione (stereo aerea o mono satellitare), digitalizzazione di carte e vettorizzazione (semi-manuale).
- Eventuale riproiezione (es. piane-geografiche)
- Inserimento attributi
- Correzione geometrica
- Correzione topologica
- Generalizzazione (per generare altre scale)

Geoprocessing Vettoriale

- Interrogazioni sugli attributi
- Interrogazioni di misura (semi geografiche)
- Interrogazioni geografiche
- Interrogazioni Statistiche
- Algebra di aree
- Analisi delle reti
- Tipi di strumenti software

Interrogazione su attributi

- Operazione non geografica: deriva dai database classici
- Filtro secondo vincoli sugli attributi
- Visualizzazione tematica (stile collegato agli attributi, a scala di colore, a classi)
- Join (collegamento di più tabelle dati secondo chiavi comuni)

Interrogazioni di misura

- Interrogazioni semi-geografiche
- Si interrogano i dati geografici, ma il risultato è un dato numerico o testuale
- Es:
 - Misura di lunghezza di oggetti lineare
 - Misura di aree o perimetri di oggetti areali
 - Misura di distanze

Interrogazioni Spaziali

- Queste interrogazioni comportano l'utilizzo di relazioni spaziali.
- Esempi tipici sono:
 - Tocca
 - E' entro la distanza di x metri da
 - Contiene (o contiene strettamente)
 - Si sovrappone a
 - E' tangente a
- Con questo tipo di interrogazioni posso analizzare l'interazione di più classi di oggetti (o di oggetti diversi della stessa classe).

Interrogazioni Statistiche

- Corrispondono alle funzioni aggreganti della basi di dati
- Con questo tipo di operazione posso ricavare dati totali numerici, es:
 - Somme totali di valori
 - Somme parziali raggruppate per criteri (es. somma popolazione centri abitati raggruppata per comuni)
 - Minimi, Massimi, Medie
- Inoltre posso ottenere dati aggregati geografici:
 - Massimi ingombri (MBR, BBOX, Insiemi Convessi)
 - Aree totali, lunghezze totali, baricentri

Algebra di Aree

- Queste operazioni generano nuovi oggetti geometrici, con operazioni algebriche, a partire da quelli esistenti:
 - Intersezione di aree
 - Differenza di aree
 - Unioni totali o parziali secondo regole (merge selettivo)
 - Buffer zone



Analisi di Reti

- In presenza di una topologia a grafo (reti) posso eseguire la Network Analisi
- Divenuta famosa con i navigatori satellitare e Google indicazioni stradali:
 - Calcolo Percorso più breve fra A e B (in senso di tempo, in senso di distanza), eventualmente con vincoli (evita pedaggi)
 - Reintradamento in caso di eventi accidentali (cantieri, incidenti)
 - Ottimizzazione di trasporti e magazzino
 - Tracking in tempo reale del trasporto (es. corrieri espressi)
 - Etc.

Tipologia degli Strumenti Software

- Command Line: es. tool di grass, GDAL
- Dialoghi con parametri: es. Tool di ArcGIS
- Interfacce GIS: Geomedia, ArcGIS (es. editing)
- Modelli automatizzabili:
 - A basso livello: programmazione standard o con linguaggi di scripting (Python di Arcgis, Visual Basic di Geomedia)
 - Ad alto livello: Model Builder di ArcGIS!