

Paolo Mogorovich

Sistemi Informativi Territoriali

Appunti dalle lezioni

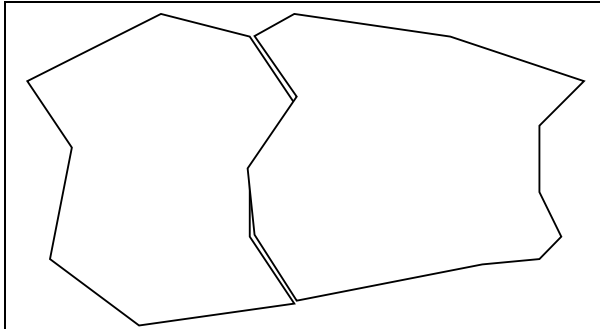
La topologia (versione provvisoria) Cod.271 - Vers.E00

- 1 - La geometria imperfetta**
- 2 - La primitiva arco**
- 3 - La topologia delle aree**
- 4 - La geometria perfetta**
- 5 - La topologia delle linee**
- 6 - I concetti base della topologia**
- 7 - Modelli topologici**
- 8 - Sistemi topologici e sistemi non topologici**
- 9 - La regola di Eulero**

1 - La geometria imperfetta

Immaginiamo la rappresentazione tramite primitive vettoriali di due oggetti territoriali di tipo areale. Supponiamo anche che gli oggetti originari (p.e. i limiti amministrativi di due Comuni adiacenti) siano

nella realtà tra di loro perfettamente adiacenti, nel senso che tra di essi non esistono né spazi vuoti né aree di sovrapposizione.



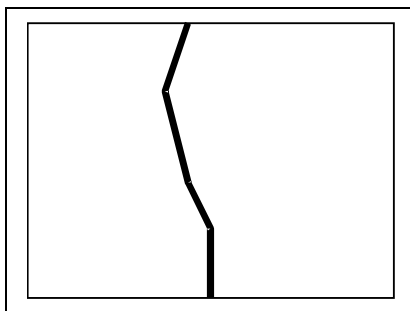
La figura evidenzia, con ovvie esagerazioni, l'incoerenza tra due aree che sono adiacenti nel mondo reale, ma non nella rappresentazione informatica. L'incoerenza tra le due aree è in genere molto piccola e spesso irrilevante in una fase di disegno in quanto inferiore allo spessore del tratto della penna; tuttavia a livello informatico questa incoerenza non è tollerata perché può portare a risultati contraddittori.

Non è detto necessariamente vero che questa adiacenza sia rispettata nel momento in cui vengono acquisite le due aree che descrivono i nostri oggetti; infatti nella fase in cui viene creata l'informazione geografica numerica, le coordinate di un vertice di un'area, per quanto acquisite con grande precisione, non saranno esattamente uguali a quelle del corrispondente vertice dell'area adiacente.

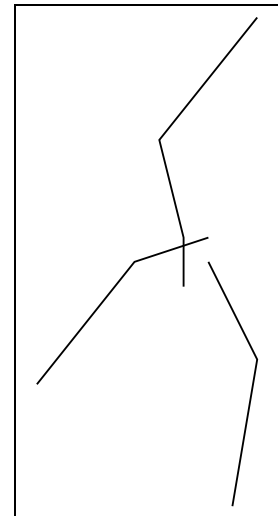
Lo stesso problema può valere per una struttura lineare. In questo caso l'acquisizione di elementi lineari che convergono in un vertice sarà tale che, ai fini di un'analisi puramente visiva, la coincidenza dei punti terminali apparirà perfetta; tuttavia a livello numerico i valori dei due punti saranno lievemente diversi.

L'incoerenza tra le due aree nel primo

caso (o tra le tre linee nel secondo) è in genere molto piccola e spesso irrilevante in una fase di disegno in quanto inferiore allo spessore del tratto della penna; tuttavia a livello informatico questa incoerenza non è tollerata perché può portare a risultati contraddittori. Ad esempio un oggetto molto piccolo (un pozzo, un traliccio, ecc.) vicino al confine potrebbe risultare, a seguito di una procedura di calcolo, come appartenente ad ambedue i Comuni, o a nessuno dei due.



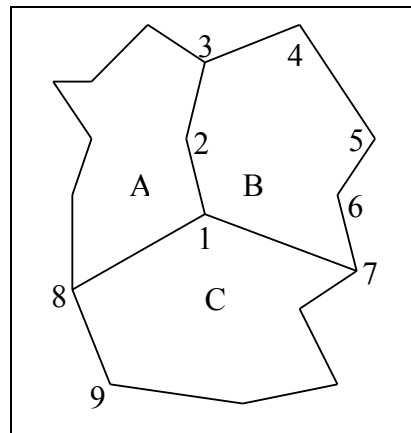
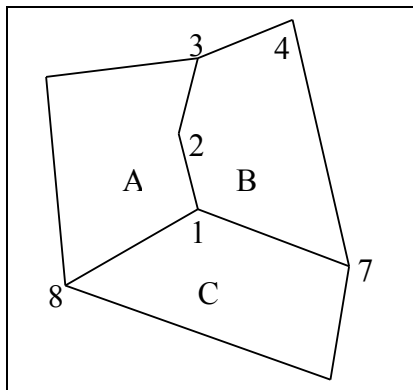
Un altro elemento di incertezza si ha quando si rappresenta il bordo di un'area e di questo se ne vede soltanto una parte: non abbiamo alcuna informazione per sapere se la parte interna dell'area si trovi a destra o a sinistra della linea. Soltanto l'aggiunta di ulteriori elementi grafici è in grado di farci capire dove è la parte interna dell'area.



2 - La primitiva arco

Consideriamo un insieme di tre aree formate da un certo numero di segmenti collegati tramite vertici. Cerchiamo di semplificare le linee, togliendo i punti ritenuti meno importanti.

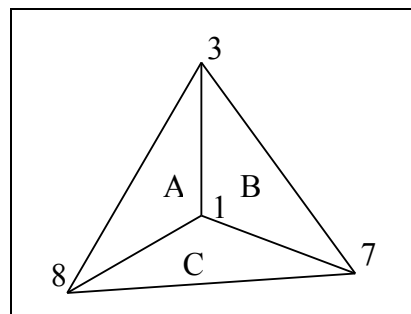
È abbastanza intuitivo cominciare a semplificare togliendo i vertici 2, 5, 6, 9 e altri fino a raggiungere il risultato della seconda figura. Un'ulteriore semplificazione può portare a eliminare il punto 4 ed anche altri, fino ad arrivare ad una struttura molto semplice.



A questo punto si percepisce chiaramente che non sono eliminabili altri punti, in quanto si modificherebbe non più la geometria, ma anche una parte “strutturale” della figura. L'eliminazione del

punto 7, ad esempio, causerebbe l'eliminazione delle aree B e C, e non soltanto una loro semplificazione geometrica.

Da quanto detto sopra, si può ipotizzare che, in una struttura di aree, i singoli elementi areali possano essere rappresentati non dalla poligonale chiusa che circonda l'area, ma da una combinazione di elementi più semplici. L'unità informativa elementare non è più la poligonale, ma una sua parte, cioè la linea che divide due aree: con riferimento alle figure l'unità informativa elementare potrebbe essere la linea che unisce i punti 3 e 7; nelle figure tale linea divide l'area B dall'esterno del disegno, e questa caratteristica rimane costante qualunque sia la ricchezza del dettaglio geometrico.



Questa unità informativa elementare si chiama “arco” e può essere formata da un numero imprecisato, maggiore o uguale a uno, di segmenti. Due o più archi convergono in un punto detto nodo.

Un'area può essere definita come un insieme di archi. Se usiamo questo modo di definire le aree vengono eliminate le incongruenze sui bordi notate in precedenza, in quanto la linea che definisce un'area è la stessa linea che definisce l'area contigua.

Se immaginiamo che l'arco sia una primitiva dotata, come il punto, la linea e l'area, di attributi, e se consideriamo come attributi di un arco il “quale area c'è a destra” e il “quale area c'è a sinistra” si vede che l'arco rispetta il concetto di omogeneità di attributi, mentre il perimetro no. La definizione di questi attributi presuppone, tramite il concetto di destra e sinistra, di definire un verso di percorrenza dell'arco.

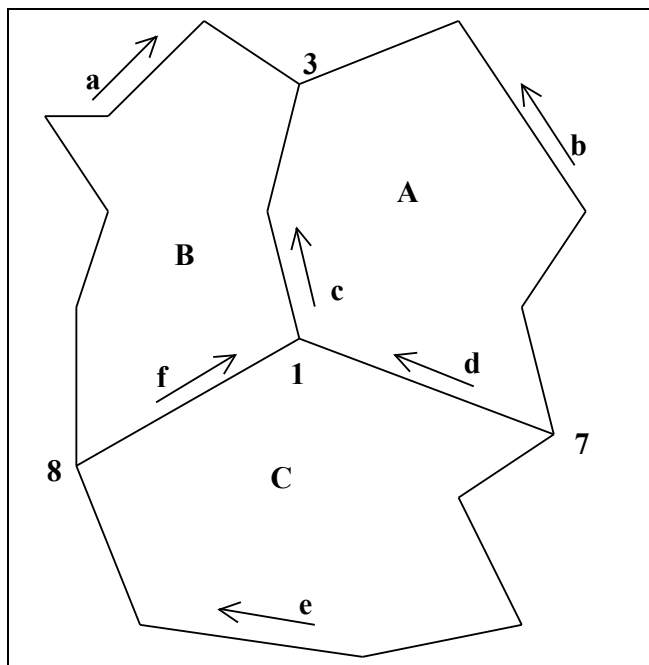
La conoscenza di questi attributi ci permetterebbe inoltre di risolvere l'altra difficoltà notata, consistente nel non sapere a priori, dato un tratto di bordo, se l'area descritta è a destra o a sinistra.

Abbiamo quindi definito una nuova primitiva avente la geometria della linea, con l'aggiunta di un “verso” e con due attributi. Cercheremo di usare questa primitiva per descrivere le primitive geometriche area e linea in modo diverso da quello semplice delle primitive geometriche vettoriali.

3 - La topologia delle aree

Perché un'area possa essere codificata correttamente come un insieme di archi, questi devono essere orientati e ci deve essere una convenzione sul rapporto tra il verso della linea e il "pieno" dell'area; adottiamo la regola che "il dentro è a destra", cioè camminando sulla linea nel suo verso troviamo l'area alla nostra destra (l'adozione della regola simmetrica non avrebbe cambiato alcunché). Con questa ipotesi, data una mappa costituita da un insieme di aree adiacenti, è possibile costruire una tabella (tabella degli archi) che associ ad ogni arco gli attributi Sin e Des, cioè i codici delle aree che si trovano rispettivamente a sinistra e a destra.

<i>Id. arco</i>	<i>Area Sin.</i>	<i>Area Des.</i>
a	--	B
b	A	--
c	B	A
d	C	A
e	--	C
f	B	C



La tabella che definisce le aree si può ricavare immediatamente dalla tabella degli archi: un'area è definita dagli archi che hanno il codice dell'area o nel campo "Area Sin" o nel campo "Area Des". Se

infatti un arco ha tra i suoi attributi "Area Sin" o "Area Des", il codice di un'area, vuol dire che esso contribuisce a separare tale area dalle altre e quindi necessariamente definisce il bordo dell'area.

Se il codice dell'area è nel campo "Area Des", l'arco va preso con suo verso, altrimenti col verso opposto.

<i>Id. Area</i>	<i>Composizione</i>
A	-b +c +d
B	+a -c -f
C	-d +e +f

La tabella degli archi si presta a considerazioni interessanti. Operando su di essa infatti è possibile eseguire rapidamente alcune operazioni, come ad esempio:

Contiguità: due aree sono contigue se esiste almeno un arco che ha nei campi Area_Sin e Area_Des i codici delle due aree (con riferimento alla figura, le aree A e B sono contigue in quanto l'arco "c" contiene i loro codici nei campi Area_Sin e Area_Des).

Unione: l'area unione di due aree è quella formata dagli archi che hanno il codice di una sola delle due aree nei campi Area_Sin e Area_Des; vanno esclusi sia gli archi che hanno nei campi Area_Sin e Area_Des codici diversi, sia quelli che hanno nei campi Area_Sin e Area_Des i codici delle due aree contemporaneamente (l'unione delle aree A e B è formata dagli archi +a -b +d -f; sono esclusi dalla combinazione sia l'arco "e" che non contiene nei campi Area_Sin e Area_Des né il codice A né il B, sia l'arco "c" che invece li contiene entrambi). L'operazione di Unione è facilmente estendibile al caso di più di due aree.

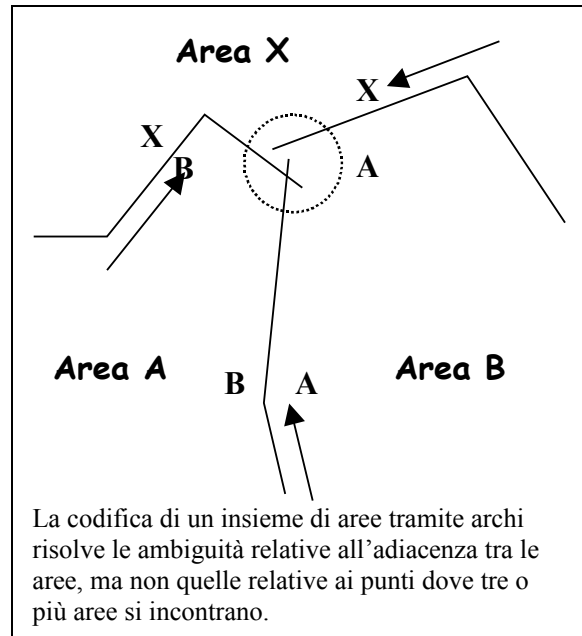
Altre operazioni che coinvolgono l'area nel suo insieme possono essere eseguite, mantenendo una struttura dati del tipo "ad archi", cioè senza ricomporre tra loro più archi per ottenere la codifica esplicita dell'area. Sono esempi il calcolo del perimetro, il calcolo della superficie, la preparazione della campitura.

4 - La geometria perfetta

L'ipotesi di costruire un'area come insieme di archi risolve il problema dell'ambiguità al bordo tra due aree. Con la soluzione adottata il bordo è unico e ciascun punto, nelle vicinanze del confine tra due aree, apparterrà inevitabilmente a una, e ad una sola, delle due aree.

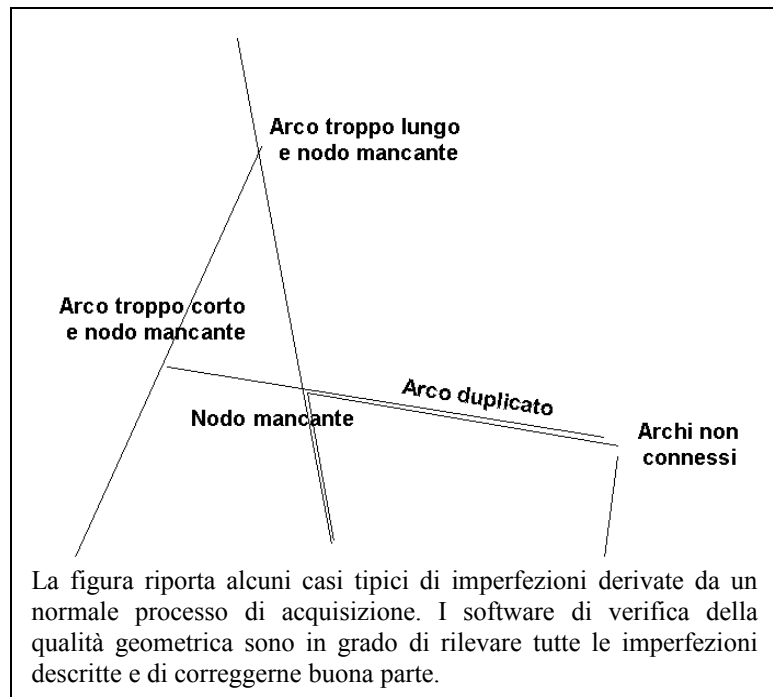
Non abbiamo però ancora risolto il problema relativo alle linee (la convergenza esatta di più linee nello stesso punto) e lo stesso problema vale nel caso delle aree. Infatti, una volta definite le aree come insiemi di archi, dobbiamo imporre l'ulteriore condizione che gli archi siano connessi con esattezza geometrica ad altri archi.

Far collassare gli estremi di un arco verso un unico punto è un'operazione puramente geometrica. Essa viene svolta dall'operatore in fase di digitalizzazione, per esempio utilizzando la funzione "snap", almeno dove ciò è possibile, o in una fase successiva, in cui opportune funzioni proprie del software si incaricano di analizzare tutti gli archi esistenti e di verificare che l'estremo di ciascun arco coincida "esattamente" (cioè abbia coordinate numericamente uguali) con l'estremo di almeno un altro arco.

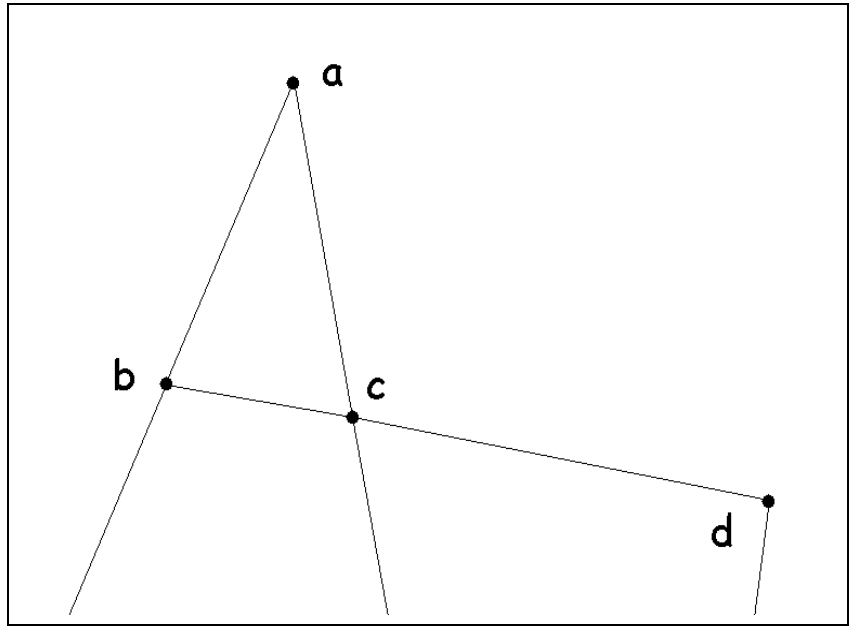


Quando la situazione precedente non è verificata, il software analizza se, in un intorno predefinito e in genere abbastanza piccolo, esistono altri estremi di archi; se così è gli estremi degli archi vengono fatti collassare automaticamente nello stesso punto che assume il ruolo di nodo. Se la correzione non è possibile il software segnala una situazione su cui occorre intervenire.

Questa operazione di pulizia analizza diversi casi, oltre a quello considerato, ed è in genere integrata a funzioni interattive di editing grafico. Le correzioni eseguite dal software, poiché avvengono all'interno di tolleranze predefinite, non inficiano la qualità geometrica dei dati, in quanto le tolleranze stesse sono state definite in modo opportuno.



L'operazione di pulizia che abbiamo descritto opera su archi, prescindendo dal fatto che tali archi descrivano un insieme di aree o di linee. Supponendo di pulire i dati della figura precedente avremmo alla fine una situazione come quella riprodotta a lato. Si noti che nei nodi "b" e "c" convergono tre o più archi, mentre nei nodi "a" e "d" convergono solo due archi. Se in un nodo convergono due archi, i loro codici "area_a_sin" e "area_a_des" sono necessariamente uguali e quindi i due archi possono essere riuniti in un solo arco. Sempre con riferimento alla figura, quindi, i nodi "a" e "d" sono falsi nodi, concettualmente errati anche se a volte tollerati dalle strutture informatiche.



L'unico caso in cui un nodo che unisce due soli archi è concettualmente corretto è quello in cui si ha un'area definita da un solo arco, cioè un'isola.

Il risultato finale di un'operazione di pulizia geometrica è un insieme di archi che hanno le seguenti caratteristiche:

- tutte le intersezioni tra archi sono dei nodi
- dato un nodo e un cerchio di raggio R_t intorno ad esso, non esistono altri nodi né altri archi all'interno di quell'area
- dato un arco, non esiste alcun vertice di nessun altro arco a distanza minore di R_t da esso.

R_t è una quantità collegata alla risoluzione geometrica dei nostri dati; se due elementi sono tra loro più distanti di R_t , essi sono due elementi distinti; se essi sono tra loro distanti meno di R_t , non devono essere considerati come elementi distinti.

Quando i dati geometrici rispettano queste condizioni la loro geometria è perfetta (da un punto di vista topologico) ed è corretto utilizzarli per costruire strutture topologiche.

5 - La topologia delle linee

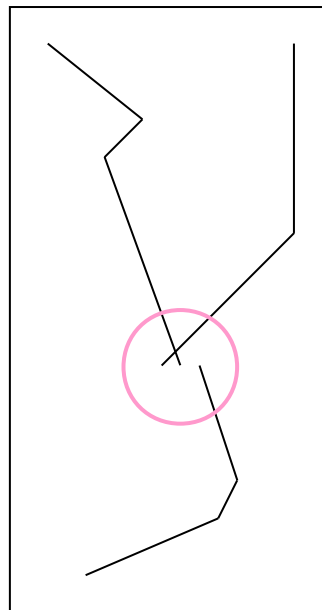
Il problema dell'adiacenza tra due aree (elementi a due dimensioni) è stato risolto introducendo i concetti di arco e nodo, e utilizzando gli archi, accorpati in insiemi, per definire le aree. Esiste un problema analogo nel caso di elementi a una dimensione, cioè le linee.

In questo caso esiste un'unica ambiguità, consistente nella non perfetta coincidenza degli estremi di archi convergenti. Immaginiamo una rete stradale, codificata in modo lineare. Gli incroci tra strade potrebbero apparire incongruenti, come in figura. Il problema della connessione tra linee diverse (anche in questo caso dette archi, pur se non dividono aree) avviene ipotizzando che non esistano diversi punti terminali per elementi convergenti, bensì uno solo, detto anche qui nodo, in modo analogo al procedimento usato per l'adiacenza tra aree.

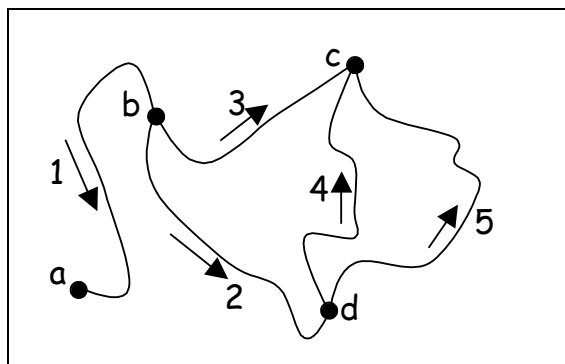
Da un punto di vista puramente geometrico, i programmi di pulizia descritti sono in grado di creare dati geometricamente perfetti; si noti che, a differenza del caso delle aree, un nodo isolato è in una struttura di linee pienamente legittimo (p.e. una strada a fondo chiuso).

Una volta ottenuti dati geometricamente perfetti, è possibile, come nel caso delle aree, costruire delle strutture tabellari che esplicitino relazioni tra elementi geometrici.

Si riporta un esempio di un insieme di linee e di una tabella che esprime le relazioni topologiche tra archi e nodi; si noti che la struttura topologica proposta richiede l'orientazione dell'arco, altrimenti gli attributi "archi uscenti" e "archi entranti" non sarebbero definibili. Nel caso delle strade, la tabella potrebbe essere utilizzata per la scelta del percorso ottimale tra due punti; l'operazione avviene solo sulla tabella, senza scomodare la più complessa parte geografica.



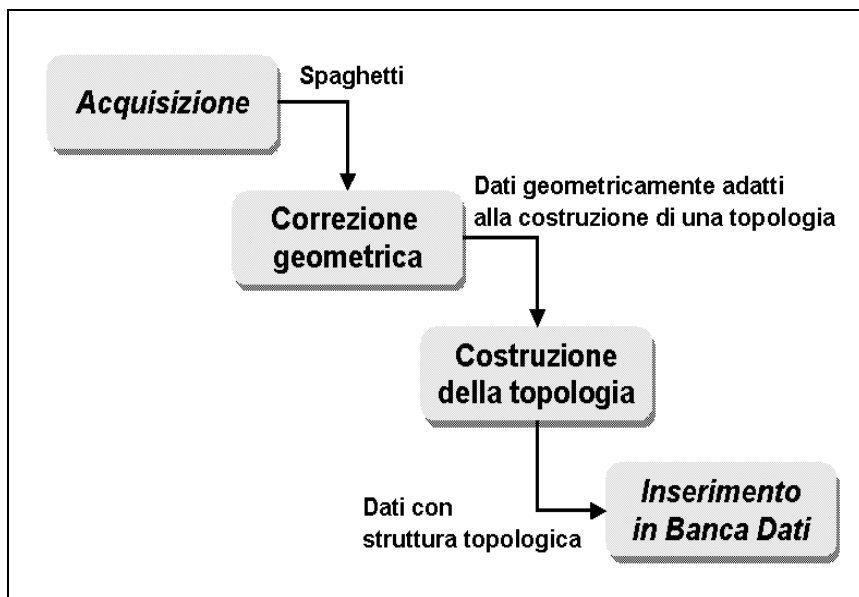
Nodo	Archi uscenti	Archi entranti
A		1
B	1, 2, 3	
C		3, 4, 5
D	4, 5	2



6 - I concetti base della topologia

La topologia studia relazioni spaziali tra oggetti geometrici; nella tecnologia GIS la topologia suggerisce l'esplicitazione di una serie di relazioni spaziali tra primitive geometriche; tali relazioni si potrebbero anche calcolare direttamente dalla geometria, ma sarebbe molto più costoso. Nello spazio della topologia, invece, una volta che le relazioni sono state esplicitate tramite tabelle, alcune operazioni sono più semplici, in quanto simili a quelle che si fanno sugli attributi.

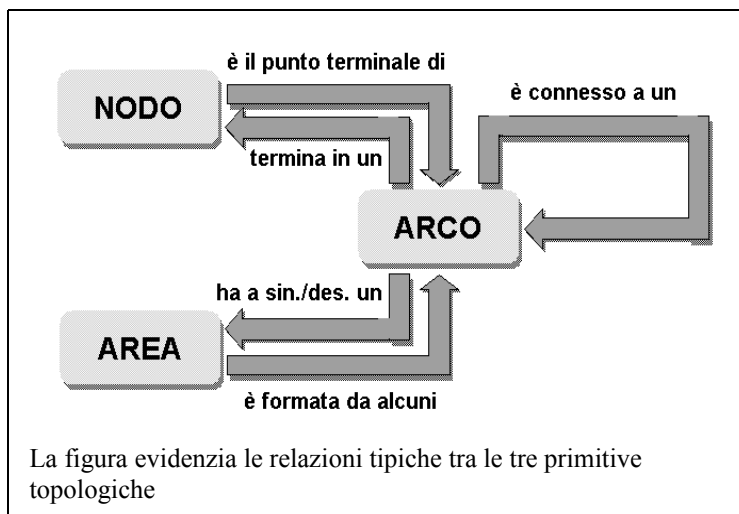
Se il dato è geometricamente perfetto (si usa dire anche “topologicamente consistente”) è possibile creare senza errori la sovrastuttura topologica; pertanto la creazione della topologia si basa su una prima fase di verifica e correzione dei dati da un punto di vista geometrico e successivamente nella creazione delle relazioni topologiche.



La topologia è quindi una sovrastuttura ridondante, che:

- garantisce la coerenza dei dati (da non confondersi con la qualità dei dati)
- dà maggiore efficienza ad alcune elaborazioni

Essa tuttavia rappresenta un costo nella preparazione dei dati, nell'importazione dei dati da altre fonti e nell'aggiornamento.



Il processo di acquisizione dati, tra le varie elaborazioni, è quello che beneficia di più di un approccio topologico. È infatti possibile organizzare il processo di acquisizione in modo da velocizzare alcune operazioni e avere contemporaneamente un controllo sui possibili errori. Questo è vero soprattutto per quanto riguarda l'acquisizione tramite digitalizzatore.

Le primitive topologiche arco e nodo trovano la loro corrispondenza geometrica nelle primitive linea e punto. Nel caso

delle aree, la primitiva areale nel mondo topologico viene chiamata “face”, ma in italiano si usa dire ancora “area”.

7 - Modelli topologici

La costruzione della topologia, cioè la costruzione di tabelle che esplicitano relazioni tra archi, nodi e aree, è tutt'altro che univoca. Infatti le relazioni definibili tra archi, nodi e aree dipendono dal fenomeno del mondo reale che viene modellato (p.e. se i dati che trattiamo definiscono una rete stradale, non vi saranno ovviamente relazioni che riguardano le aree); inoltre, a parità di modello topologico, i vari software GIS implementano le relazioni in modo diverso.

I modelli topologici più comuni sono:

- Full planar graph: è la struttura topologica adatta a descrivere elementi areali contigui e non sovrapposti, tipicamente limiti amministrativi o mappe catastali.
- Planar graph linear network: è adatto a descrivere un insieme di linee interconnesse tra loro, senza la presenza di aree, col vincolo che tutte le intersezioni sono nodi; il caso tipico è quello di una rete stradale o ferroviaria ma senza ponti o sottopassi.
- Non planar graph linear network: è adatto a descrivere un insieme di linee interconnesse tra loro, senza la presenza di aree e senza il vincolo che tutte le intersezioni siano nodi. In pratica due linee si possono intersecare geometricamente, ma non strutturalmente, come nel caso di un'autostrada e di una strada secondaria che la attraversa su un ponte.
- Spaghetti Questa struttura indica un insieme di dati codificati e memorizzati senza alcun vincolo. Esso può riguardare sia dati che non hanno mai avuto una struttura topologica e quindi geometricamente imperfetti, ma può anche riguardare il caso interessante di dati che hanno avuto una struttura topologica la quale, senza modifiche alla geometria, è stata tolta. In questo caso i dati si trovano in una situazione di “correttezza” geometrica in modo tale che una struttura topologica può essere facilmente costruita.

Il Full planar graph è il modello più restrittivo e per lui valgono le seguenti regole, definite e adottate, oltre 20 anni fa, dall'U.S. Census Bureau,

- ogni arco ha due nodi
- ogni arco divide due aree
- ogni area è circondata da archi e nodi
- ogni nodo è circondato da aree e archi
- tutte le intersezioni sono nodi

Se immaginiamo un contesto lineare, per esempio una rete stradale, vediamo che:

- la seconda e la terza regola non hanno senso
- la quarta va modificata
- la quinta regola non è sempre vera, come accade ad esempio nel caso dei sottopassi e dei cavalcavia.

						• ogni arco ha due nodi
						• ogni arco divide due aree
						• ogni area è circondata da archi e nodi
						• ogni nodo è circondato da aree e archi
						• tutte le intersezioni sono nodi
★	★	★	★	★	★	• Full planar graph
★			▲	★		• Planar graph linear network
★			▲			• Non planar graph linear network
★	▲	★	▲			• Non planar graph linear network with surfaces
★						• Spaghetti

La tabella riporta le relazioni esistenti tra alcuni modelli topologici e regole che definiscono le relazioni tra le primitive topologiche. L'asterisco indica il rispetto della regola, il triangolo un rispetto parziale.

È possibile mettere in relazione queste regole con i modelli topologici, come riportato in tabella. L'asterisco indica il rispetto della regola, il triangolo un rispetto parziale, come è evidente, ad esempio nel caso del "Planar graph linear network" dove la regola "ogni nodo è circondato da aree e archi" vale nella forma ridotta "ogni nodo è circondato da archi".

8 - Sistemi topologici e sistemi non topologici

Il fatto che, a parità di modello topologico, i vari software GIS implementino le relazioni in modo diverso pone il problema del trasferimento di dati tra sistemi diversi. Nel passaggio di dati da un sistema ad un altro, l'unica possibilità è che la topologia venga "smontata" dal sistema di uscita e "ricostruita" dal sistema che riceve i dati secondo le diverse regole di quest'ultimo. Vi sarà quindi un formato intermedio in cui i dati si muovono da un sistema all'altro, e in questa fase i dati non hanno una struttura topologica (non esistono le tabelle di relazione tra primitive topologiche), ma mantengono le caratteristiche di qualità geometrica, in pratica restano "geometricamente perfetti". Il modello spaghetti è quello che, per definizione, si presta a questa operazione.

Un caso comune di "smontaggio" della topologia è quello secondo cui si ricostruiscono le primitive geometriche a partire dalle primitive topologiche.

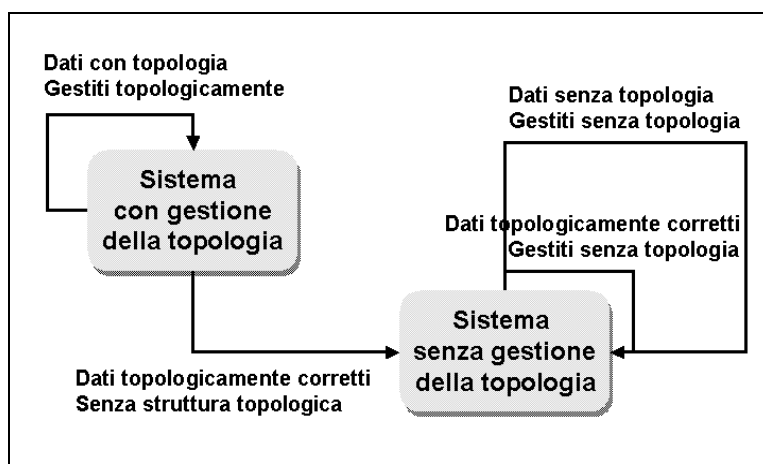
Supponiamo che all'interno di un sistema A esista una struttura topologica, per uno strato informativo areale, secondo la quale gli elementi geometrici sono descritti tramite archi e tabelle. In genere il sistema A sarà dotato di un software di trasformazione che, dagli archi e dalle tabelle, ricostruisce gli elementi areali. Questi dati non sono gestiti da una topologia, ma sono topologicamente corretti (o geometricamente perfetti), nel senso che una linea che divide l'area 1 dall'area 2 è esattamente la stessa sia quando compare nella primitiva areale dell'area 1 che in quella dell'area 2.

Quando il sistema B, dotato di un certo modo di gestire la topologia, riceve questi dati, ricostruisce gli archi, dividendo a pezzi le varie aree, e l'arco di divisione tra l'area 1 e l'area 2 estratto dall'area 1 è numericamente identico a quello estratto dall'area 2. Non vi è quindi alcuna fase di "pulizia" e i dati, inseriti nel sistema B con una nuova struttura topologica, saranno geometricamente identici ai dati che erano presenti nel sistema A.

È comune il caso in cui dati topologicamente corretti, ma senza struttura topologica, vengano portati in un sistema software di livello basso (C), che non gestisce strutture topologiche, ma che semplicemente tratta le primitive geometriche così come sono.

Il sistema C, per sua natura, se è dedicato ad acquisizione di dati, crea archivi dove non è possibile che i dati siano topologicamente corretti.

Tuttavia, nel caso in cui riceva da un altro sistema dati topologicamente corretti, siamo nella fortunata situazione di avere dati geometricamente perfetti all'interno di un software che non gestisce la topologia (dati topologicamente corretti gestiti senza topologia). Questa fortunata situazione dura finché i dati vengono utilizzati solo in operazioni di "lettura", senza modifiche. Se i dati vengono modificati è molto probabile che essi non siano più geometricamente perfetti, e il sistema C non è in grado di attivare un processo di costruzione topologica che li renderebbe di nuovo geometricamente perfetti (anche se in realtà alcuni artifici permettono a volte di non deteriorare la qualità dei dati).



Questo caso è estremamente comune. Si pensi ai server che distribuiscono dati, in genere topologicamente corretti, a client dove sono attivi software di livello medio basso (tipicamente dei “desktop GIS”), che non gestiscono la topologia.

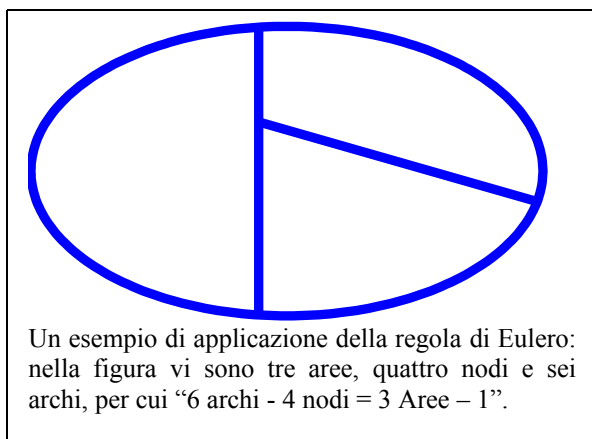
9 - La regola di Eulero

La regola di Eulero è una regola molto generale che definisce una relazione tra archi, nodi e aree di un grafo. Secondo tale regola

$$a - n = A - 1$$

dove a è il numero di archi, n il numero dei nodi, A il numero delle Aree.

La definizione di arco e nodo è più generale di quella usata parlando di topologia, nel senso che un punto isolato rispetta la regola (0 archi - 1 nodo = 0 Aree - 1) e analogamente una linea isolata (1 arco - 2 nodi = 0 Aree - 1). Nel caso di un'unica area, essa si intende formata da un arco che si chiude su se stesso in un nodo; in tale caso si ha che 1 arco - 1 nodo = 1 Area - 1 .



La regola di Eulero, inoltre, tollera i falsi nodi. Infatti, data una qualunque figura dove esistano archi, se inseriamo in un arco un falso nodo avremo un nodo in più e un arco in più, per cui la differenza $a-n$ resterà invariata.

La formula di Eulero vale nel caso in cui tutti gli elementi grafici si tocchino (insieme connesso). Se ciò non accade, vale una forma più generale della formula di Eulero:

$$a - n = A - nic$$

dove nic indica il numero degli insiemi connessi. Nella versione breve, come si è visto, nic è uguale a 1.

Nella figura a fianco è compresa l'area della figura precedente, una linea e due punti. In totale 7 archi, 8 nodi e 3 Aree. Se consideriamo che in questo caso $Nic = 4$, la formula è ancora verificata.

