ANNO ACCADEMICO 2005-2006
SISTEMI INFORMATIVI GEOGRAFICI

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

Sistemi Informativi Territoriali

2. La rappresentazione di informazioni spaziali concetti di base

ALBERTO BELUSSI

NOVEMBRE 2005

Informazione spaziale
Un Modello per lo schema dei dati: GEO-ER

Per la descrizione della componente spaziale del dato geografico diversi sono i possibili approcci:

- Approccio basato sugli insiemi di punti dello spazio Euclideo a due o tre dimensioni.
- Approccio vettoriale basato sulla geometria lineare dello spazio Euclideo a due o tre dimensioni
- Approccio basato sulle relazioni spaziali qualitative: relazioni topologiche, relazioni basate sulla direzione o sulla distanza.

Approccio basato sugli insiemi di punti dello spazio Euclideo a due dimensioni

(M. Worboys, "GIS: a computing perspective")

Spazio di riferimento: Piano Euclideo ($\Re^2$)
Oggetti: insiemi di punti

Topologia del Piano Euclideo

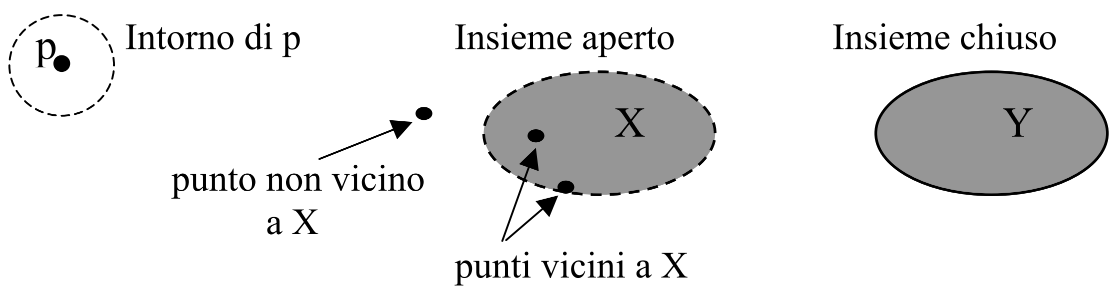
Definizioni:

Disco aperto: è un insieme di punti di $\Re^2$ limitato da una circonferenza (circonferenza esclusa).

Intorno di un punto $p \in \Re^2$: è qualsiasi disco aperto contenente p.

Dato un insieme di punti $X \subset \Re^2$: si introducono le seguenti definizioni:

- Un punto $p \in \Re^2$ è vicino a (near) X se ogni intorno di p contiene punti di X.
- X è aperto se per ogni punto $p \in X$ esiste un intorno di p contenuto in X.
- X è chiuso se contiene tutti i punti vicini a X.
- Chiusura di X (X^-): è l'unione di X con l'insieme di tutti i suoi punti vicini.
- Interior di X (X°): è l'insieme di tutti i punti di X che non sono vicini al complemento di X ($X' = \Re^2 - X$).
- Boundary di X (∂X): è l'insieme di tutti i punti che sono vicini sia a X che a X' (si dimostra che: $\partial X = X^- - X^\circ$).



Topologia del Piano Euclideo (continua)

Definizioni:

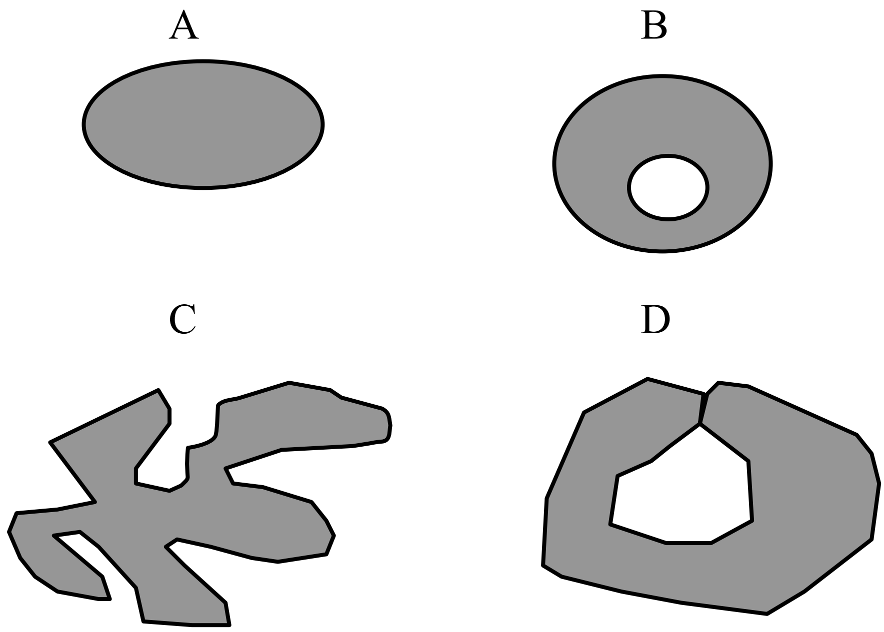
Un insieme di punti $X \subset \Re^2$ si dice un <u>insieme chiuso regolare</u> se e solo se: $X^{\circ\,-} = X$

Un insieme di punti $X \subset \Re^2$ si dice <u>(path-)connesso</u> se per ogni coppia di punti appartenenti a X esiste un cammino (path) completamente contenuto in X che li congiunge.
Cammino (definizione intuitiva): curva senza interruzioni.

<u>Omeomorfismo</u>: è una biiezione definita su $\Re^2$ tale che trasforma ogni intorno del dominio in un intorno del codominio:

$$h: \Re^2 \rightarrow \Re^2$$

Due insiemi X e Y sono topologicamente equivalenti se esiste un omeomorfismo che applicato a X produce Y.
Un omeomorfismo equivale ad una trasformazione "rubber sheet" (foglio di gomma) che stira e distorce il piano (foglio) senza però produrre pieghe o strappi.



Possibili definizioni di tipi geometrici nell'approccio basato sugli insiemi di punti.

***Definizione* di Poligono**

Un **poligono** è l'unione di:

- insiemi chiusi regolari e connessi di $\Re^2$ (estensione maggiore di zero);
- linee (archi semplici): insiemi di $\Re^2$ omeomorfi a un segmento di retta (quindi connessi);
- punti isolati: punti di $\Re^2$ non contenuti nei primi due tipi di insiemi;

tale che valga la seguente proprietà:

- ogni poligono contiene almeno un insieme chiuso regolare e connesso con estensione maggiore di zero,
- le linee e i punti isolati sono opzionali.

Gli insiemi chiusi regolari e connessi possono eventualmente contenere buchi, in tal caso il complemento del poligono non è connesso.
Le linee e i punti isolati rappresentano poligoni degenerati.

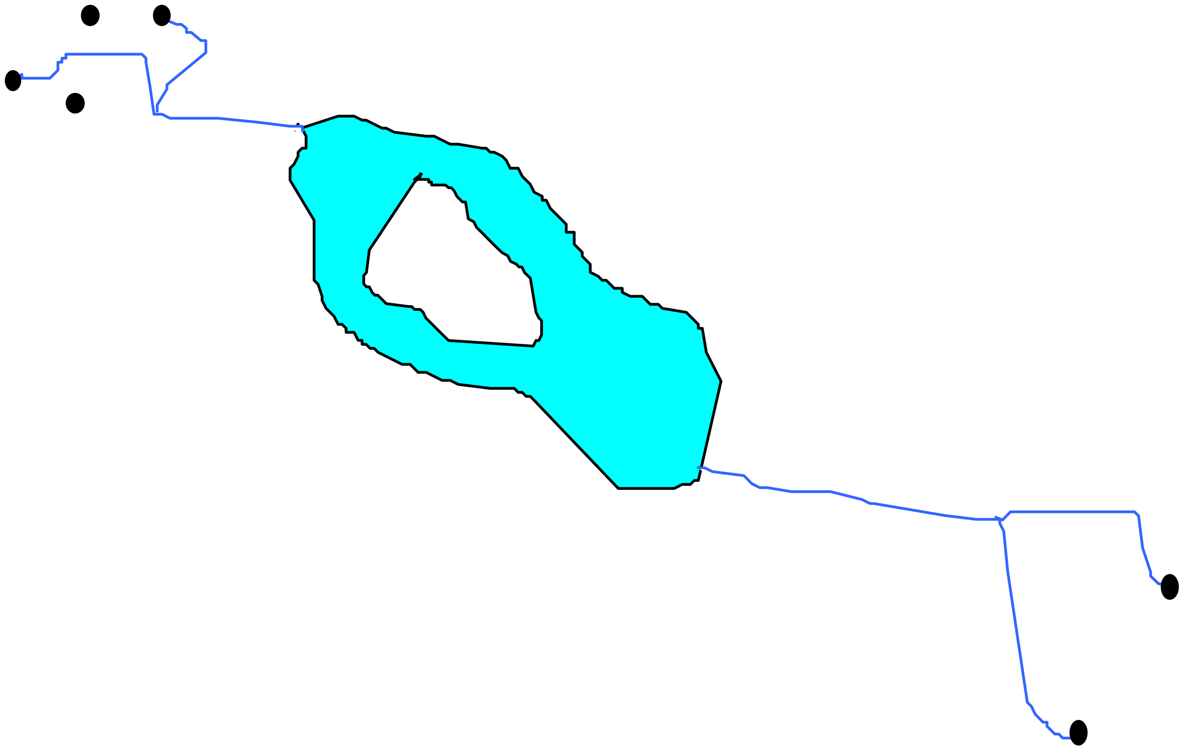
Frontiera (o boundary) di un poligono P: coincide con la definizione topologica di boundary: ∂P.

Quindi ∂P è costituito dai bordi esterni dei poligoni, compresi i bordi degli eventuali buchi, e dalle linee e punti che costituiscono poligoni degenerati.

Parte interna (o interior) di un poligono P: si ottiene per differenza con la frontiera di P: $P° = P - \partial P$

Il perimetro di un poligono P si calcola considerando anche il bordo dei buchi e contando due volte la lunghezza dei poligoni degenerati a linee.

Un esempio di **poligono**



***Definizione di* Linea**

Una **linea** è l'unione di:

- archi semplici: insiemi di $\Re^2$ omeomorfi a un segmento di retta (quindi connessi);
- punti isolati: punti di $\Re^2$ non contenuti nel primo tipo di insiemi;

tale che valga la seguente proprietà:

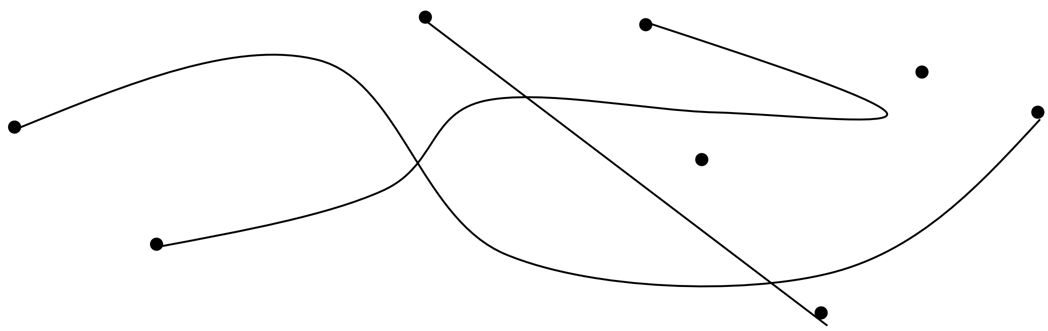
- ogni linea contiene almeno un insieme omeomorfo a un segmento di retta,
- i punti isolati sono opzionali.

I punti rappresentano curve degenerate.

<u>La frontiera (o boundary) di una linea L</u>: è costituita dagli estremi di tutti gli archi non contenuti in altri archi di L e dai punti isolati.

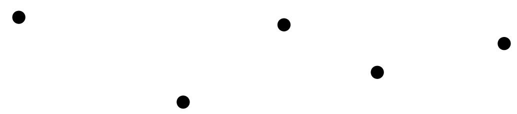
<u>La parte interna (o interior) di una linea L</u>: si deriva per differenza dalla frontiera.

La lunghezza di una linea è la somma della lunghezza di tutti gli archi semplici che compongono la linea.



Definizione di **Tipo Geometrico Punto**

E’ un qualsiasi insieme di punti isolati del piano. Per definizione la frontiera di un punto è vuota. ❑

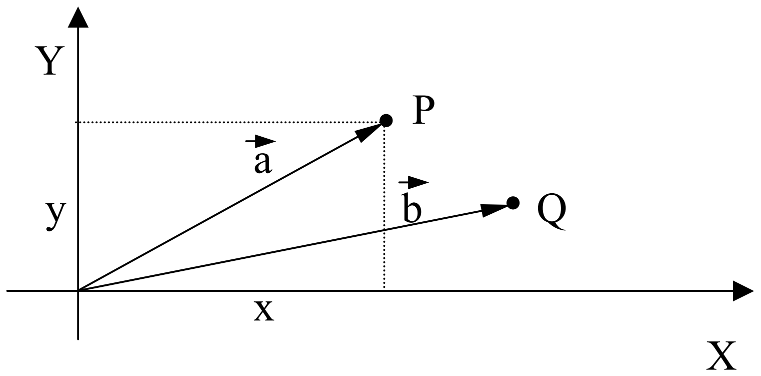


Approccio vettoriale basato sulla geometria lineare dello spazio Euclideo a due dimensioni

(M. Worboys, "GIS: a computing perspective")

Spazio di riferimento: Piano Euclideo ($\Re^2$)

Punto: P = (x,y)



$\vec{a}$ è il *vettore* che individua il punto P in $\Re^2$.

Distanza tra due punti P e Q individuati dai vettori $\vec{a} = (x_a,y_a)$ e $\vec{b} = (x_b,y_b)$:

$$\overline{PQ} = \|\vec{a} - \vec{b}\| = \sqrt{(x_b-x_a)^2 + (y_b-y_a)^2}$$

Operazioni sui vettori:

$\vec{a} + \vec{b} = \vec{c}$	$\vec{c} = (x_a+x_b,\ y_a+y_b)$
$\vec{a} - \vec{b} = \vec{d}$	$\vec{d} = (x_a-x_b,\ y_a-y_b)$
$\lambda\vec{a} = \vec{e}$	$\vec{e} = (\lambda x_a,\ \lambda y_a)$

<u>Segmento di retta</u> per i punti P e Q individuati dai vettori $\vec{b}$ e $\vec{a}$:

$$s = \{\lambda\vec{a} + (1-\lambda)\vec{b} \mid \lambda \in [0,1]\}$$

chiamiamo *estremi del segmento* i punti: $s.E_1 = Q$ (λ=1) e $s.E_0 = P$ (λ=0)

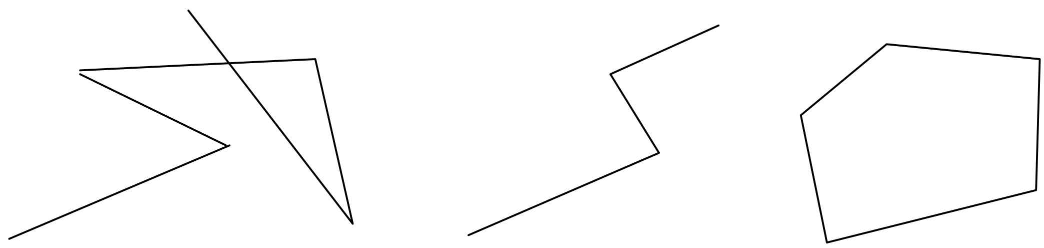
<u>Spezzata (Polyline)</u>: è un insieme di segmenti di retta $sp = \{s_1,\dots,s_n\}$ tali che ogni estremo di segmento è condiviso da <u>esattamente due segmenti</u>, eccetto al massimo due estremi detti *estremi della spezzata*.

Data la spezzata $sp = \{s_1,\dots,s_n\}$, se vale la seguente condizione:

$$(\forall s_i, s_j \in sp)(i \neq j \Rightarrow (s_i \cap s_j = \varnothing \vee s_i.E_0 = s_j.E_0 \vee s_i.E_0 = s_j.E_1 \vee s_i.E_1 = s_j.E_0 \vee s_i.E_1 = s_j.E_1))$$

(la spezzata non si autointerseca), allora la spezzata si dice *spezzata semplice*.

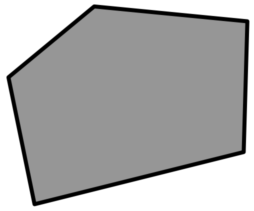
Se una spezzata non ha estremi si dice *spezzata chiusa*.



spezzata spezzata semplice spezzata chiusa

Ogni spezzata *sp* è orientata in quanto è definito quale è il suo primo estremo ($sp.E_0$) e quale il secondo ($sp.E_1$)

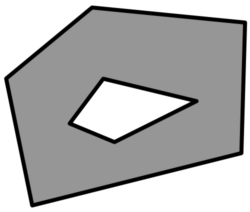
Poligono semplice (Simple Polygon): data una spezzata chiusa *sp*, è la porzione di $\Re^2$ delimitata da *sp*. *sp* fa parte del poligono (il poligono è un insieme di punti chiuso).



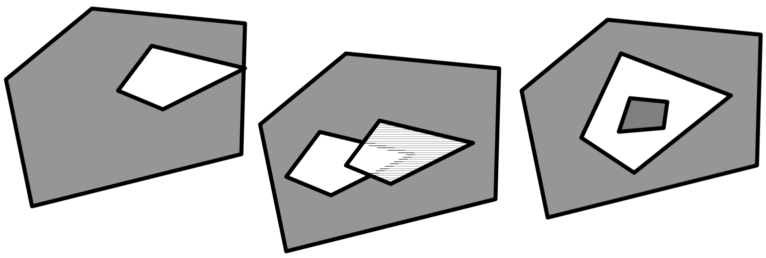
poligono semplice

Poligono bucato (Polygon): data una spezzata *sp* e un insieme di spezzate chiuse H=$\{sp_1, ..., sp_n\}$, è la porzione di $\Re^2$ delimitata da *sp* alla quale siano tolte le porzioni di piano delimitate da ciascuna $sp_i \in$ H. La spezzata *sp* e tutte le spezzate di H sono parte del poligono (il poligono è un insieme di punti chiuso), inoltre valgono le seguenti condizioni:

- tutte le spezzate di H sono contenute nel poligono semplice delimitato da *sp*.
- tutte le spezzate di H non intersecano *sp*. (*sp* interseca *sp'* se esiste almeno un segmento *s* di *sp* e un segmento *s'* di *sp'* tali che $s \cap s' \neq \emptyset$)
- $\forall(sp_i, sp_j \in H)(i \neq j \Rightarrow ((sp_i$ non interseca $sp_j) \wedge$ (sp_i non è contenuta nel poligono delimitato da sp_j))



poligono bucato



poligoni non ammessi

Relazioni nello spazio

Per descrivere proprietà spaziali è possibile, invece che descrivere la forma e la collocazione degli oggetti nello spazio, specificare le relazioni spaziali che intercorrono tra gli oggetti immersi nello spazio (ad esempio, indicando che la Provincia di Verona è adiacente alla Provincia di Vicenza si descrive in modo qualitativo una proprietà spaziale delle entità Provincia di Verona e Provincia di Vicenza).
Analizzeremo di seguito le relazioni spaziali più usate nell'ambito delle basi di dati geografiche.

Una relazione binaria definisce una proprietà che lega due valori di un determinato insieme. Nel nostro caso i valori sono geometrici e la proprietà riguarda la loro reciproca posizione nello spazio.

Esistono tre classi di relazioni spaziali binarie:

1. LE RELAZIONI TOPOLOGICHE,
2. LE RELAZIONI BASATE SULLA DIREZIONE,
3. LE RELAZIONI BASATE SULLA DISTANZA.

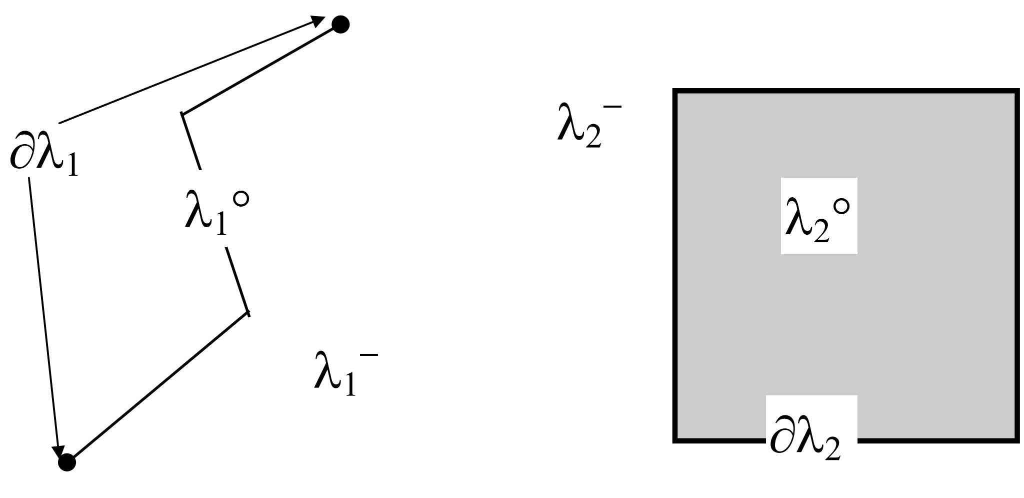
Le relazioni topologiche

Le relazioni topologiche sono definite come quelle relazioni che sono invarianti rispetto a trasformazioni topologiche (rubber sheet transformation). Esse hanno inoltre le seguenti proprietà:

- sono completamente indipendenti dalla distanza tra i due valori geometrici considerati e dall'estensione dei due valori geometrici
- sono di tipo qualitativo: descrivono "come" sono in relazione due valori geometrici senza specificare "quanto" (misura).
- fanno riferimento a concetti di alto livello, che hanno una facile corrispondenza con il linguaggio naturale.

La definizione teorica delle relazioni topologiche fa riferimento a diversi modelli. Fra questi in particolare si è affermato, nell'ambito delle basi di dati geografiche e non solo, il modello di Max Egenhofer (1991).

Tale modello considera come valori geometrici insiemi di punti chiusi e si basa sulla suddivisione dello spazio prodotta da ogni valore geometrico λ. Tale suddivisione individua: l'**interior** $\lambda°$ (parte interna), il **boundary** $\partial\lambda$ (frontiera) e l'**exterior** λ^- (esterno o complemento) di un valore geometrico λ.
Ad esempio:



In tale modello tutte le relazioni topologiche possibili tra due valori geometrici A e B si definiscono attraverso il risultato dell'intersezione tra tutte le suddivisioni dello spazio generate da A (A°, ∂A, A^-) e tutte le suddivisioni dello spazio generate da B (B°, ∂B, B^-). Formalmente:

$$R(A,B) = \begin{pmatrix} A^\circ \cap B^\circ & A^\circ \cap \partial B & A^\circ \cap B^- \\ \partial A \cap B^\circ & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^\circ & A^- \cap \partial B & A^- \cap B^- \end{pmatrix}$$

Il numero totale di relazioni è pari a 512, anche se non tutte le combinazioni corrispondono a casi reali.

Il modello è applicabile a:

- due linee nel piano cartesiano
- due poligoni nel piano cartesiano
- una linea e un poligono nel piano cartesiano
- due poligoni nello spazio a tre dimensioni
- una linea e un poliedro nello spazio a tre dimensioni
- due poliedri nello spazio a tre dimensioni

Se consideriamo il caso di due poligoni nel piano cartesiano basta considerare quattro intersezioni per avere tutti i casi significativi:

$$\begin{matrix} A^\circ \cap B^\circ & A^\circ \cap \partial B \\ \partial A \cap B^\circ & \partial A \cap \partial B \end{matrix}$$

Le relazioni topologiche

E. Clementini, P. Di Felice e P. van Oosterom hanno proposto un insieme di relazioni topologiche applicabili a tutte le tipologie (punto, linea, poligono) di valori geometrici.

Dati due valori geometrici λ_1 e λ_2 si definiscono le seguenti relazioni topologiche:

TOUCH

$$\lambda_1 \text{ TOUCH } \lambda_2 \quad \Leftrightarrow \quad (\lambda_1^\circ \cap \lambda_2^\circ = \varnothing) \wedge (\lambda_1 \cap \lambda_2 \neq \varnothing)$$

IN

$$\lambda_1 \text{ IN } \lambda_2 \quad \Leftrightarrow \quad (\lambda_1^\circ \cap \lambda_2^\circ \neq \varnothing) \wedge (\lambda_1 \cap \lambda_2 = \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$$

<: cross tra poligono e un insiemi finito di punti.

CROSS

$$\lambda_1 \text{ CROSS } \lambda_2 \quad \Leftrightarrow \quad dim(\lambda_1^\circ \cap \lambda_2^\circ) \leq (max(dim(\lambda_1),dim(\lambda_2)) - 1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$$

OVERLAP

$$\lambda_1 \text{ OVERLAP } \lambda_2 \Leftrightarrow \quad dim(\lambda_1^\circ \cap \lambda_2^\circ) = dim(\lambda_1) = dim(\lambda_2) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_1) \wedge (\lambda_1 \cap \lambda_2 \neq \lambda_2)$$

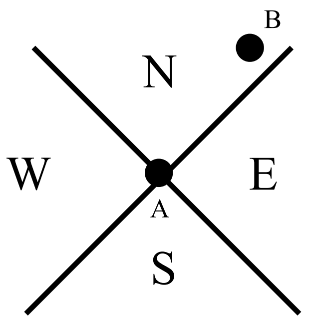
DISJOINT

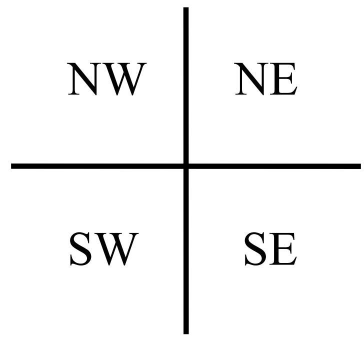
$$\lambda_1 \text{ DISJOINT } \lambda_2 \Leftrightarrow \quad (\lambda_1 \cap \lambda_2 = \varnothing)$$

dove la funzione *dim* restituisce la dimensione del risultato: 0 punto, 1 linea, 2 poligono oppure *empty* se l'intersezione è vuota.

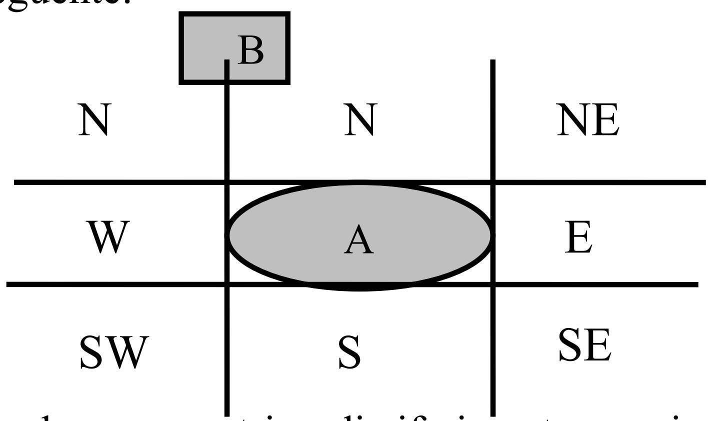
Le relazioni basate sulla direzione

Le relazioni direzionali sono definite sulla base di un sistema di riferimento direzionale che partiziona lo spazio almeno in quattro quadranti. L'origine del sistema di riferimento viene posta su uno dei due valori geometrici considerati. Intersecando i quadranti con l'altro valore geometrico si deriva la relazione tra i due valori.
Considerando valori geometrici puntiformi, le partizioni dello spazio possono essere ad esempio le seguenti:





Considerando invece valori geometrici con estensione significativa il modello da adottare deve tener conto di tale estensione e quindi può essere ad esempio il seguente:



Preso A come valore geometrico di riferimento, se si applica il modello topologico per descrivere le relazioni tra le partizioni dello spazio e l'altro valore geometrico B, si ottengono 169 relazioni.

Le relazioni basate sulla distanza

Queste relazioni richiedono la definizione nello spazio di riferimento di una funzione che calcola la distanza. Se si utilizza la distanza Euclidea allora la distanza tra due punti $P=(x_1,y_1)$ $Q=(x_2,y_2)$ è definita come segue:

$$D(P, Q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

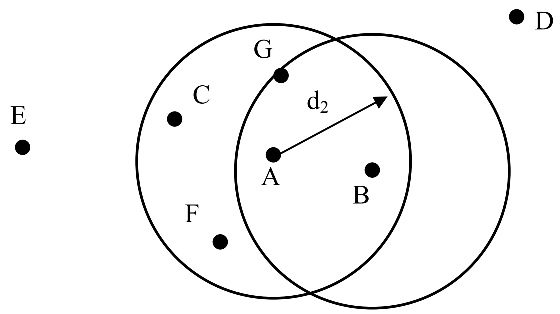
Se consideriamo invece due valori geometrici con estensione significativa, occorre ridefinire il concetto di distanza, ad esempio come segue:

$$\mathit{Dist}(A, B) = \mathit{min}(\{D(P,Q): P \in A, Q \in B\})$$

Fissata quindi una coppia di distanze (d_1,d_2) è possibile definire una relazione metrica tra due valori geometrici come segue:

$$R_{(d_1,d_2)}(A, B) \Leftrightarrow (d_1 < \mathit{Dist}(A,B) < d_2)$$

Ad esempio, considerando la relazione $R_{(0,d_2)}$:



Sono in relazione con A i punti: B, C, G, F.

Relazioni approssimate

Può essere necessario rappresentare relazioni tra valori geometrici che non derivano da misure precise di distanza, ma che si basano sui concetti astratti di vicinanza e lontananza usualmente presenti nel linguaggio naturale.

Tali relazioni sono dette **relazioni metriche approssimate**. Un insieme di relazioni metriche approssimate ha cardinalità finita, ad esempio: {same_location, very_near, near, medium, far, very_far} costituisce un insieme di relazione metriche approssimate.

E' solitamente possibile fissare un ordinamento delle relazioni, ad esempio:

same_location < very_near < near < medium < far < very_far

Quindi un insieme di relazioni metriche approssimate partiziona completamente lo spazio di riferimento. La definizione precisa di tale partizione dipende dal contesto (c'è differenza tra il "near" di chi cammina e il "near" di chi viaggia in aereo) e comunque può essere precisata attraverso degli intervalli di distanza.