Qualité des données (géographiques)

François Pinet

UR TSCF
Irstea, centre de Clermont-Ferrand

francois.pinet@irstea.fr

Qualité des données : principes de base

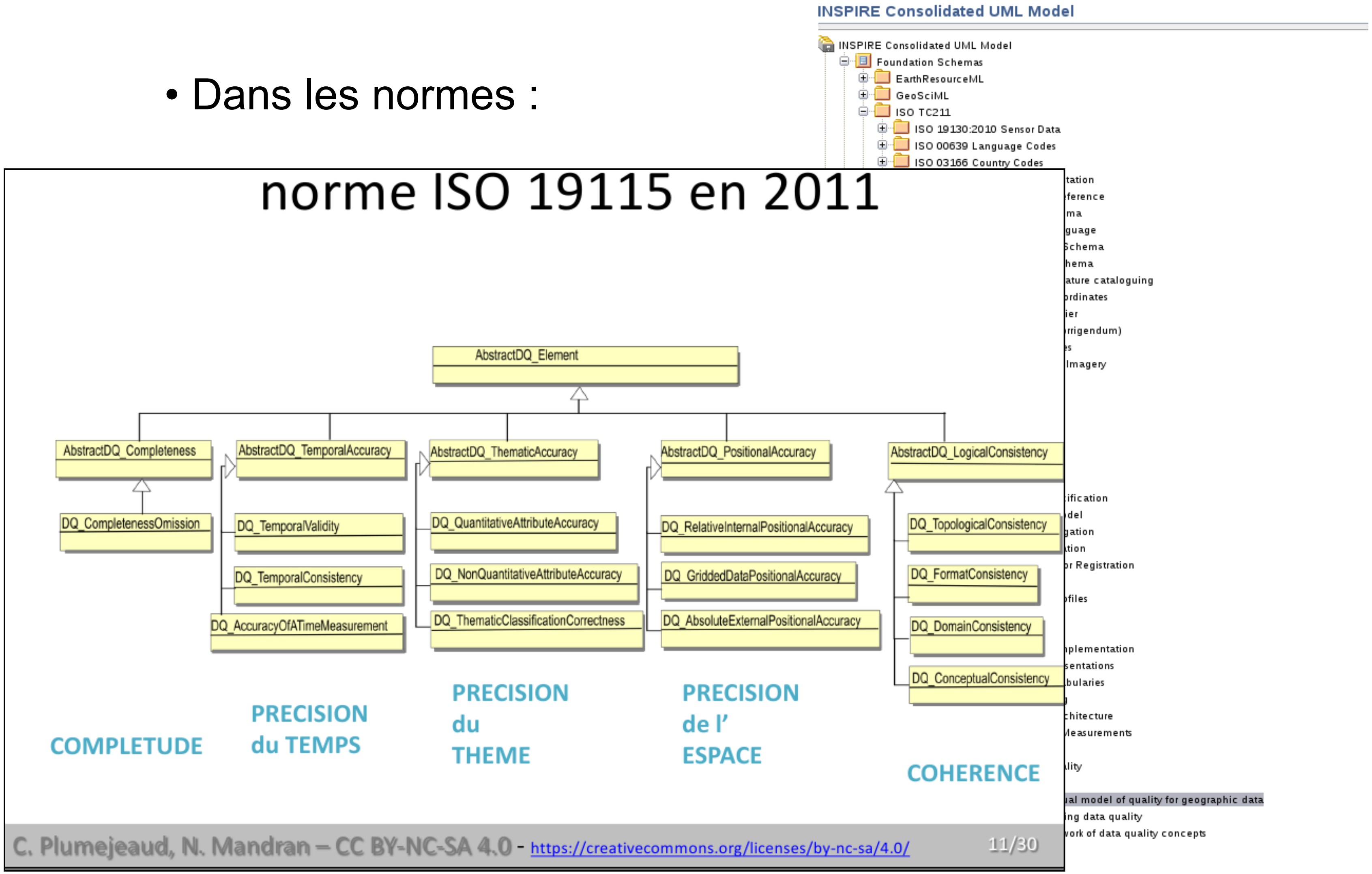
- La conformité des données à une utilisation
- Qualité interne :
  - Contrôlée par le producteur des données
  - Respect de certaines propriétés pour une utilisation générique
  - Ex. : précision métrique des données
- Qualité externe :
  - Contrôlée par l'utilisateur des données pour une application en particulier
  - « fitness for use »
  - Respect des propriétés pour une utilisation particulière
  - Ex. : les refuges de montagne sont-ils indiquées sur ma carte

Dans les normes

- Dans les normes :

Dans les normes

- Dans les normes :



La terminologie

flou, vague, imprécis,
inconsistant, ambiguë,
exhaustif, clarté, …

Qu'est-ce qu'un polygone vague ?

Un ligne floue ?

Une donnée ambiguë ?

Une entrepôt de données inadéquats ?

La terminologie

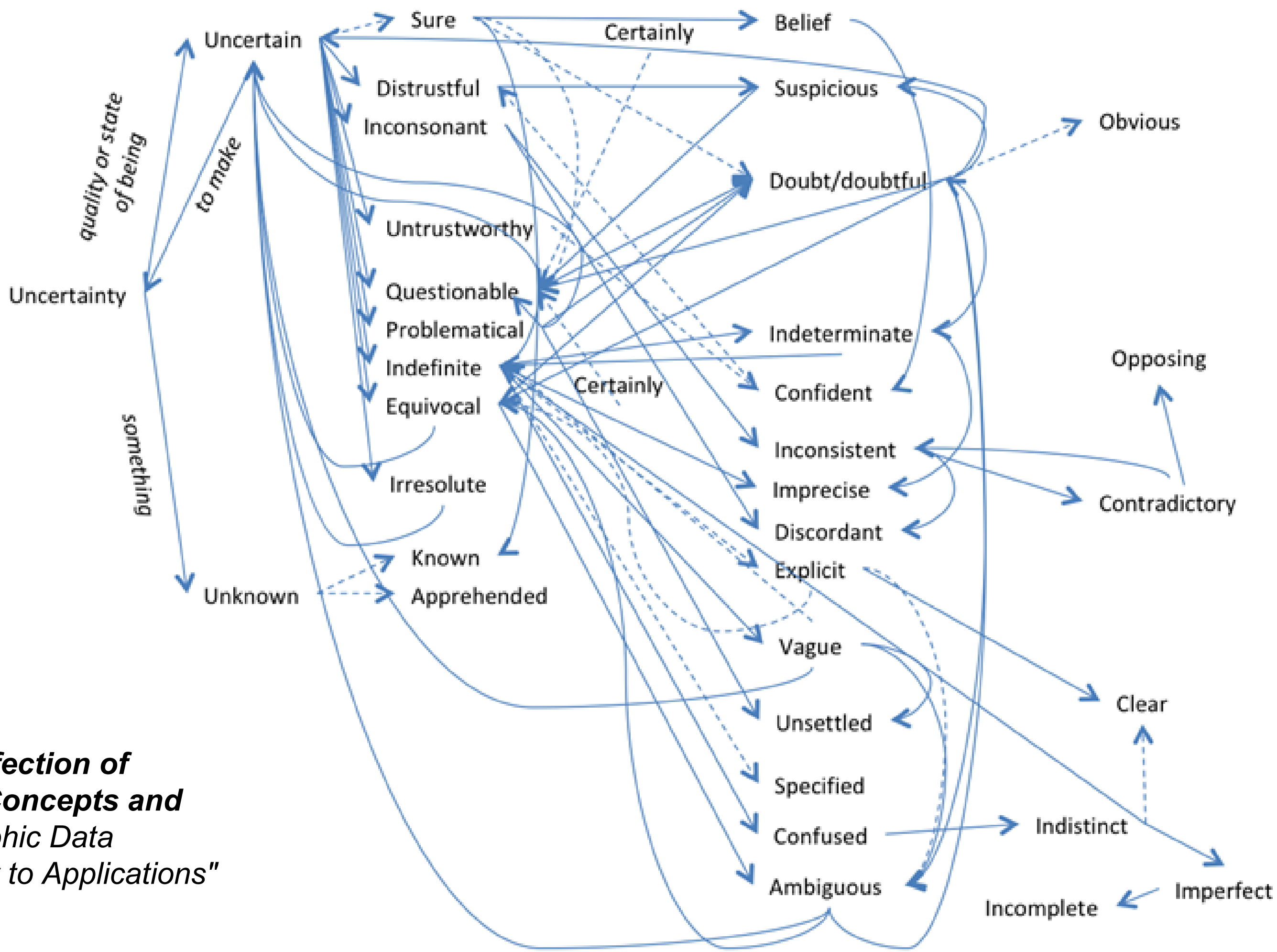
Imperfection *:*

- ***Uncertainty*** : « the difficulty to determine something ». Le problème peut venir :
  - ***Ambiguity*** : « problem to identify and categorize a given object ». Ex. *Arachnide est une espèce ou une famille d'espèces ?*
  - ***Vagueness*** : « uncertainty to find limit of objects ».
  - ***Unsuitability*** *: Ex. Les utilisateurs ne comprennent pas les données*
- ***Incorrectness*** *: « incorrect regarding the requirements » :*
  - ***Irrelevance*** *:* « not relevant to the target use »
  - ***Incompleteness***
  - ***Inconsistency*** *:* « logical contradictions »

La terminologie

- En pratique, la définition des termes est un peu plus « floue » ou changeante
- Est-ce que les scientifiques peuvent se raccrocher à des définitions génériques ?

La terminologie



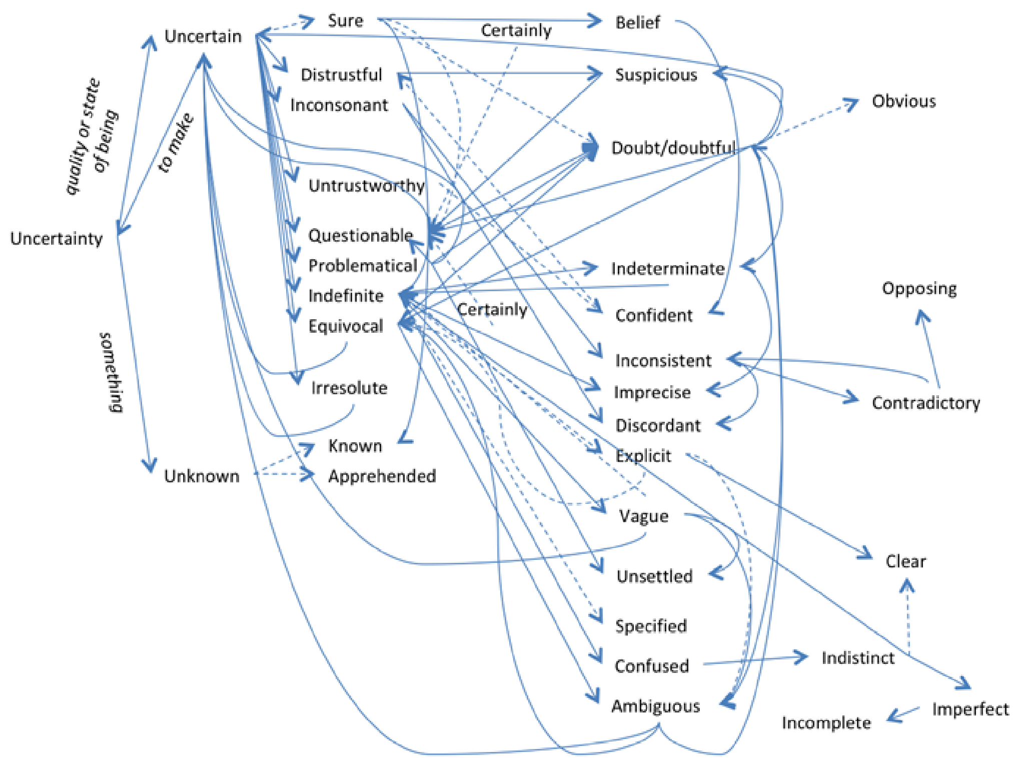
Devillers et al., 2019, ***Imperfection of Geographic Information: Concepts and Terminologies****, in "Geographic Data Imperfection 1: From Theory to Applications"*

Graphe représentant les relations entre différents mots du dictionnaire anglais Webster

La terminologie

Bilan :

- Problèmes des traductions
- Changer de sources (par ex. de dictionnaires), produit des graphes différents
- Définition bi-directionnel



Graphe représentant les relations entre différents mots du dictionnaire anglais Webster

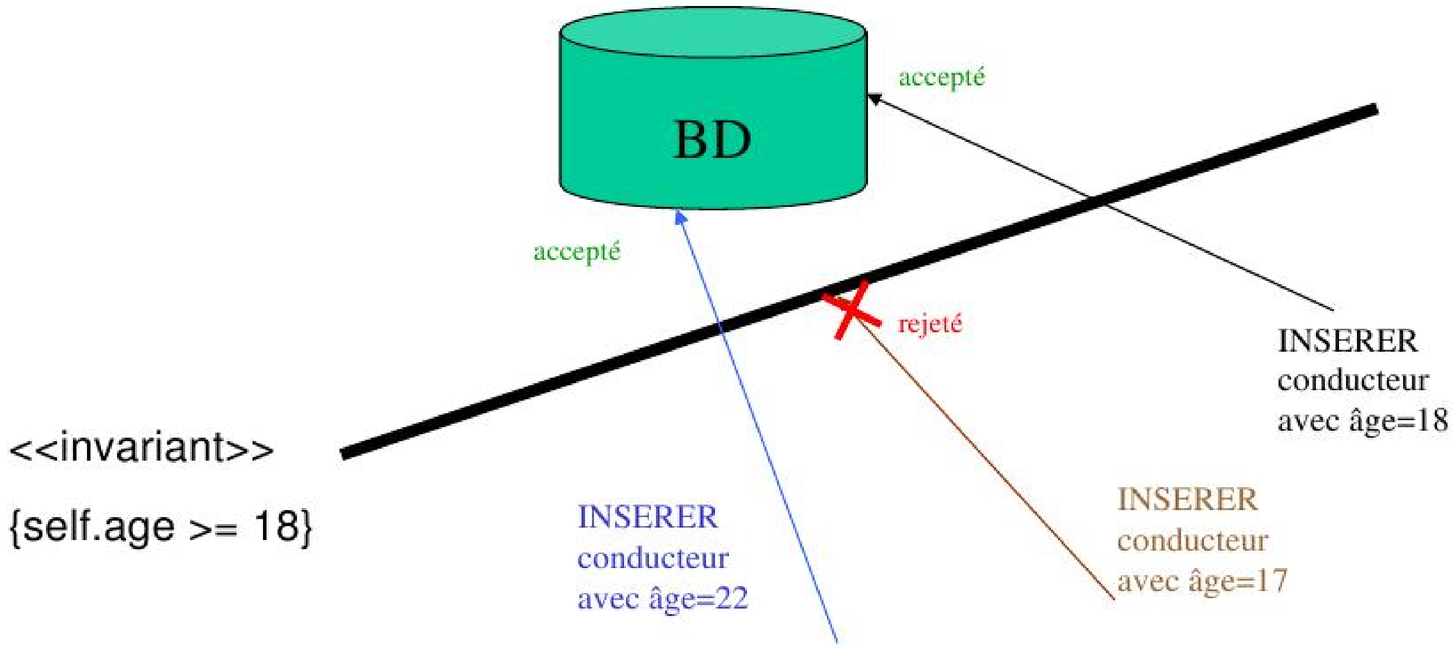
Les techniques

- Comment tester les données pour assurer leur qualité ?

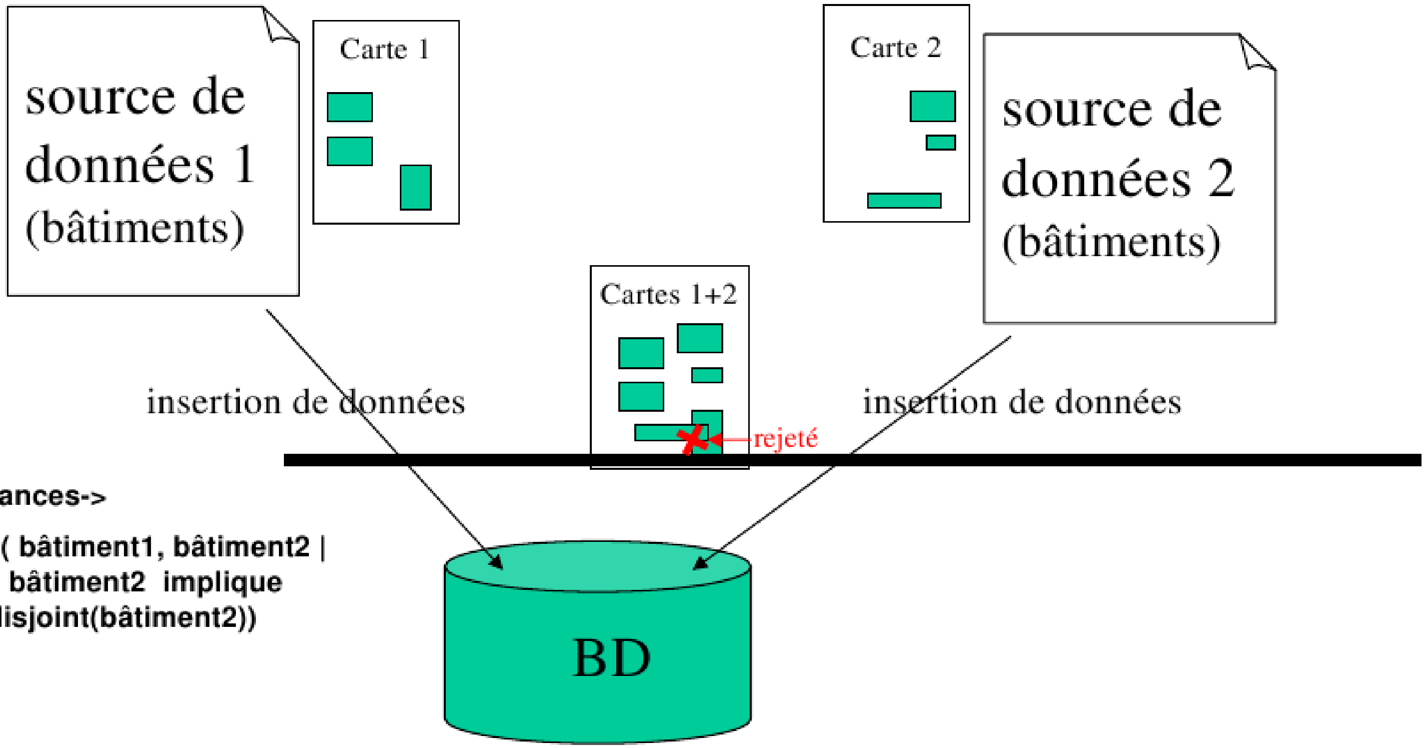
Les techniques

- Comment tester les données pour assurer leur qualité ?

  Les contraintes d’intégrité



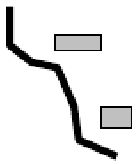
Les contraintes d'intégrité

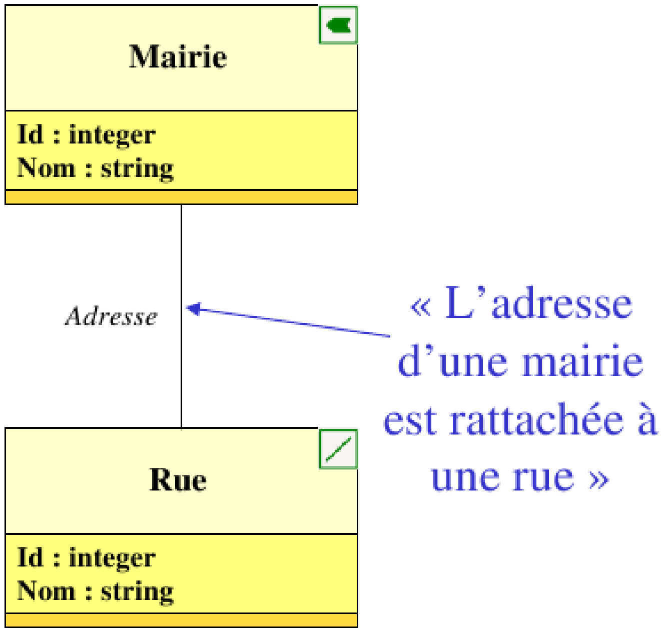


Building.Instances->

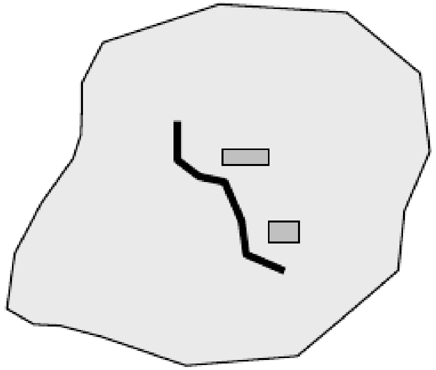
QuelqueSoit (bâtiment1, bâtiment2 | bâtiment1 <> bâtiment2 implique (bâtiment1).disjoint(bâtiment2))

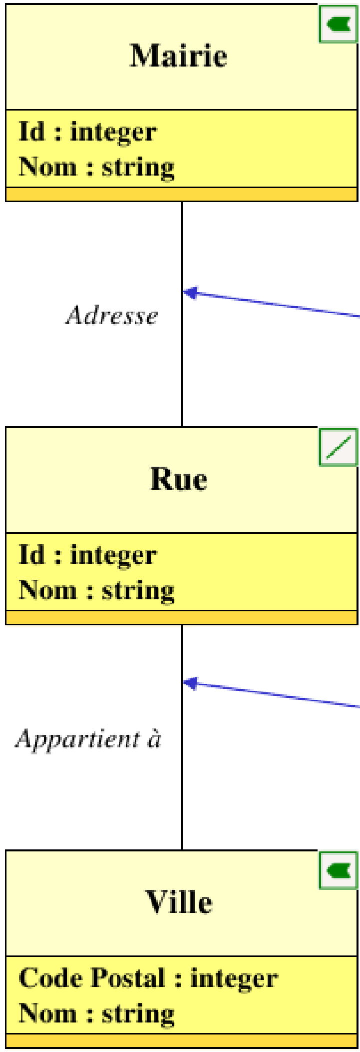
Les contraintes d'intégrité





Les contraintes d'intégrité



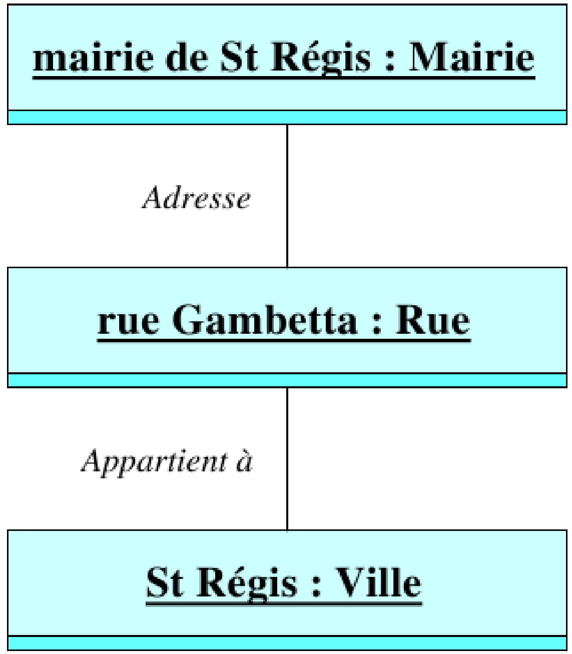


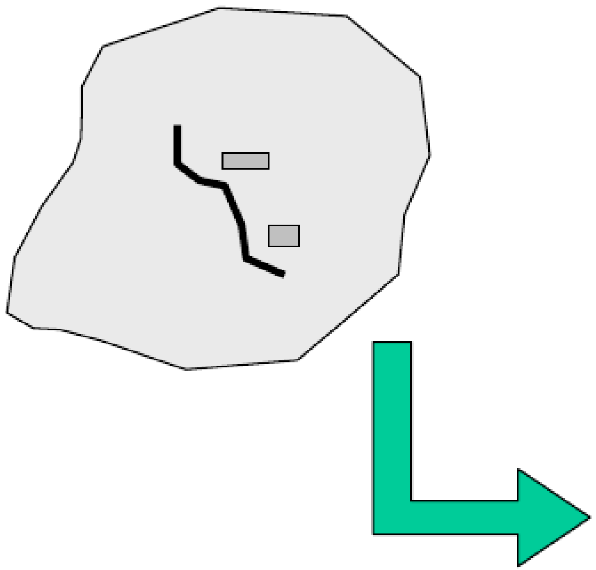
« L'adresse d'une mairie est rattachée à une rue »

« Une rue appartient à une ville »

Contrainte: «Une mairie doit être dans la ville correspondant à sa rue»

Les contraintes d'intégrité

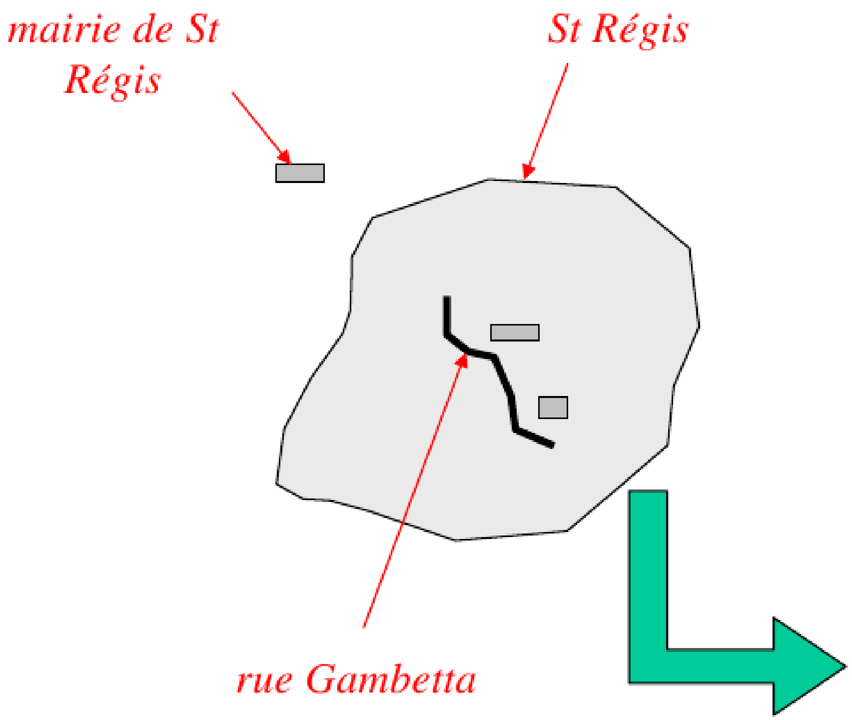


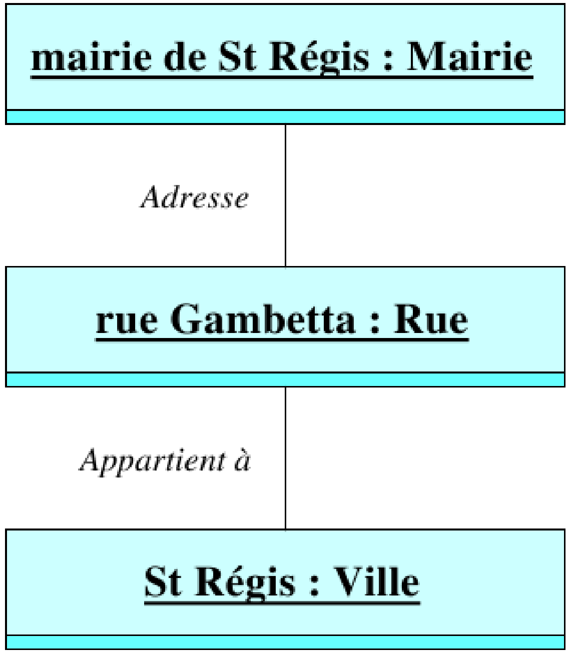


« La Mairie de St Régis se trouve dans cette ville »

Les contraintes d'intégrité

Problème de qualité des données spatiales:

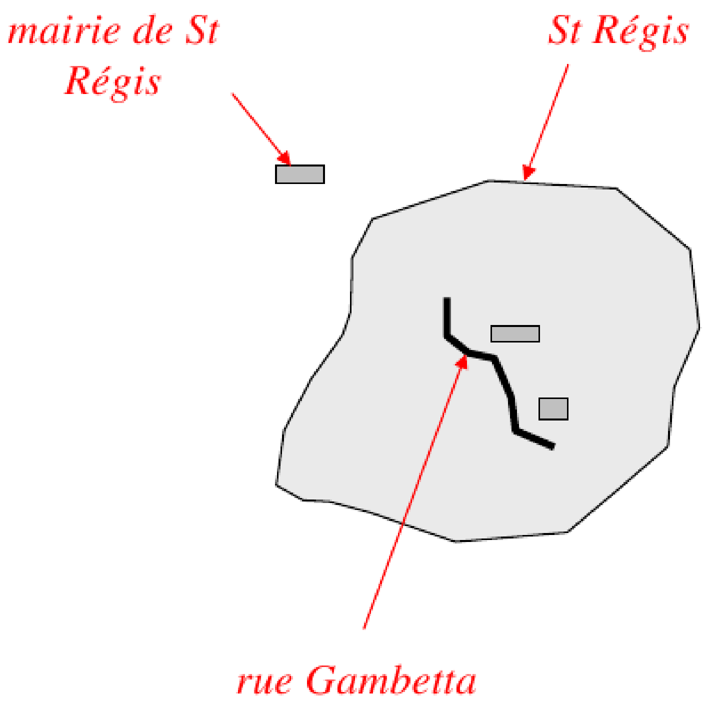




« La Mairie de St Régis se trouve dans cette ville »

Les contraintes d'intégrité

Problème de qualité des données spatiales:



Données Spatiales

Mairie (visualisation des données alphanumériques) :

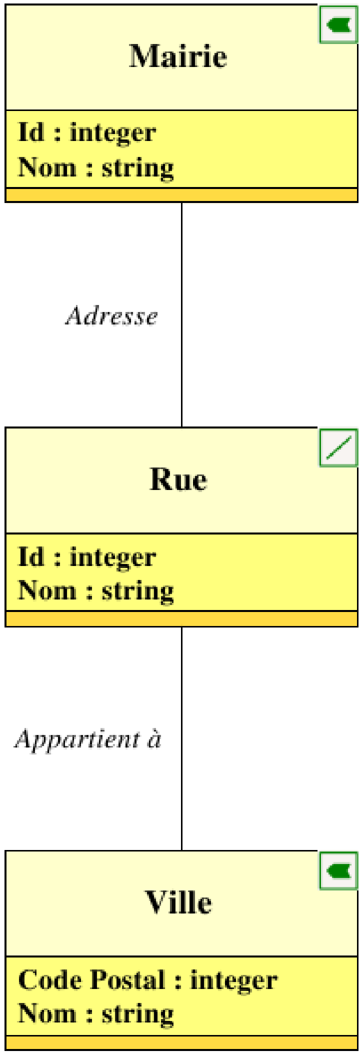
Mairie.Nom : *mairie de St Régis*

Rue.Nom : *rue Gambetta*

Ville.Nom : *St Régis*

Données Alphanumériques

Les contraintes d'intégrité



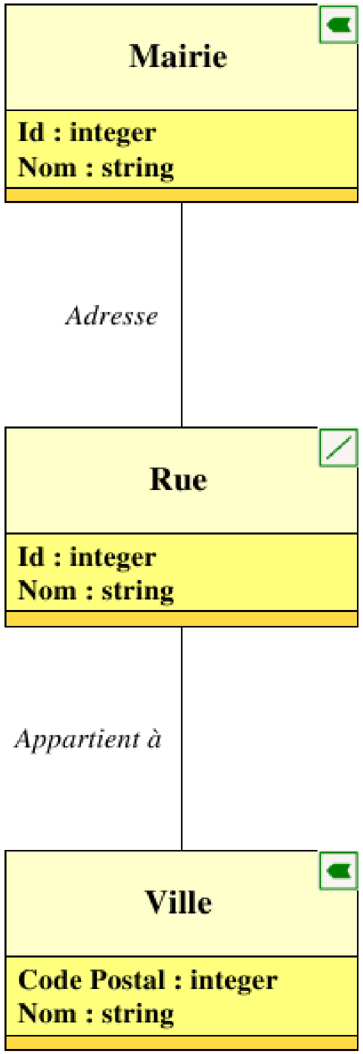
Contrainte: «Une mairie doit être dans la ville correspondant à sa rue»

=

```
context Mairie inv:
self.estDans(self.Rue.Ville)
```

une mairie

Les contraintes d'intégrité



Contrainte: «Une mairie doit être dans la ville correspondant à sa rue»

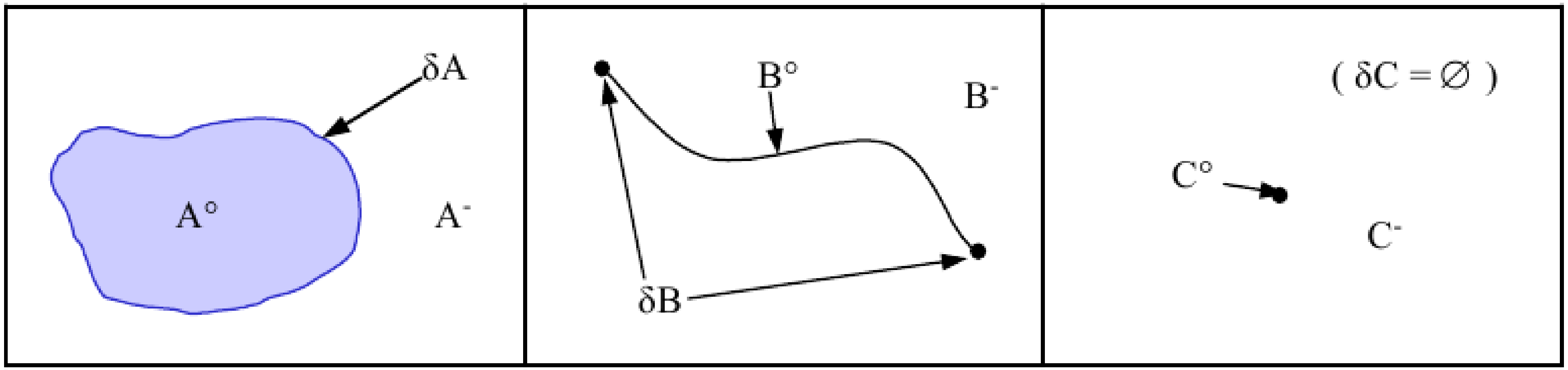
=

```
context Mairie inv:
self.estDans(self.Rue.Ville)
```

une mairie

Les relations topologiques sur objets certains

Comment définir une relation topologique

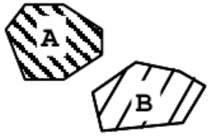


Les relations topologiques sur objets certains

$$
M = \begin{array}{c} \\ A^\circ \\ \partial A \\ A^- \end{array}
\begin{array}{ccc} B^\circ & \partial B & B^- \\
\left[\begin{array}{ccc}
A^\circ \cap B^\circ \neq \varnothing & A^\circ \cap \partial B \neq \varnothing & A^\circ \cap B^- \neq \varnothing \\
\partial A \cap B^\circ \neq \varnothing & \partial A \cap \partial B \neq \varnothing & \partial A \cap B^- \neq \varnothing \\
A^- \cap B^\circ \neq \varnothing & A^- \cap \partial B \neq \varnothing & A^- \cap B^- \neq \varnothing
\end{array}\right] & & \end{array}
$$

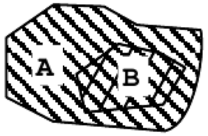
1 matrice = 1 relation topologique

(Egenhofer et al., 1990, tech. Report ; Egenhofer et al., SSD 93, 1993)



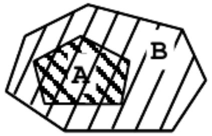
$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

⟨A, disjoint, B⟩



$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

⟨A, contains, B⟩

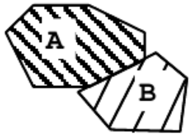


$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

⟨A, inside, B⟩

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

⟨A, equal, B⟩

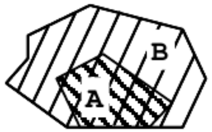


$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

⟨A, meet, B⟩

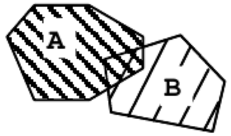
$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

⟨A, covers, B⟩



$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

⟨A, coveredBy, B⟩

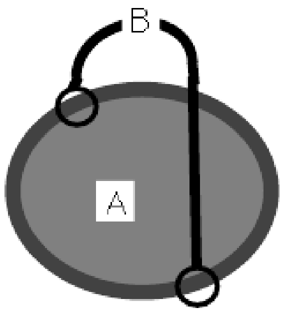


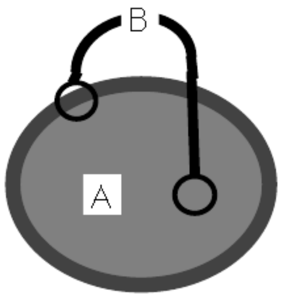
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

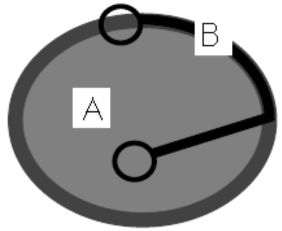
⟨A, overlap, B⟩

Les relations topologiques sur objets certains

$$
M = \begin{array}{c} \\ A^\circ \\ \partial A \\ A^- \end{array}
\begin{array}{c}
\begin{array}{ccc} \qquad B^\circ \qquad & \qquad \partial B \qquad & \qquad B^- \qquad \end{array} \\
\begin{pmatrix}
A^\circ \cap B^\circ \neq \varnothing & A^\circ \cap \partial B \neq \varnothing & A^\circ \cap B^- \neq \varnothing \\
\partial A \cap B^\circ \neq \varnothing & \partial A \cap \partial B \neq \varnothing & \partial A \cap B^- \neq \varnothing \\
A^- \cap B^\circ \neq \varnothing & A^- \cap \partial B \neq \varnothing & A^- \cap B^- \neq \varnothing
\end{pmatrix}
\end{array}
$$





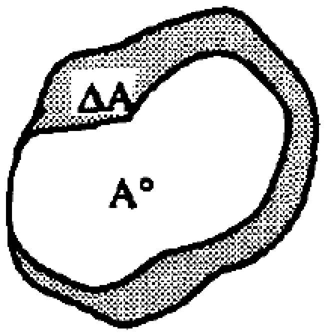


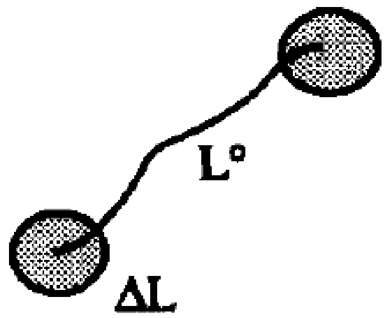
$$
\begin{array}{c} \\ A^\circ \\ \partial A \\ A^- \end{array}
\begin{array}{c}
\begin{array}{ccc} B^\circ & \partial B & B^- \end{array} \\
\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}
\end{array}
\qquad
\begin{array}{c} \\ A^\circ \\ \partial A \\ A^- \end{array}
\begin{array}{c}
\begin{array}{ccc} B^\circ & \partial B & B^- \end{array} \\
\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}
\end{array}
\qquad
\begin{array}{c} \\ A^\circ \\ \partial A \\ A^- \end{array}
\begin{array}{c}
\begin{array}{ccc} B^\circ & \partial B & B^- \end{array} \\
\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}
\end{array}
$$

(Egenhofer et al., 1990, tech. Report)

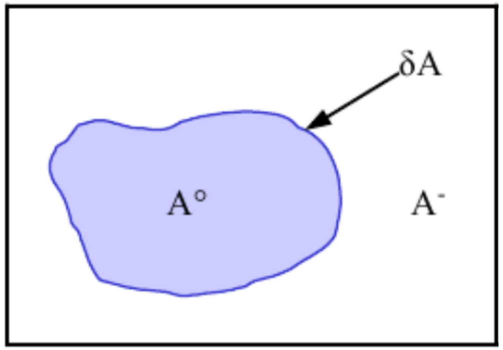
Les relations topologiques sur objets incertains

(Clementini et al., IJAR, 1997)





Les relations topologiques sur objets incertains

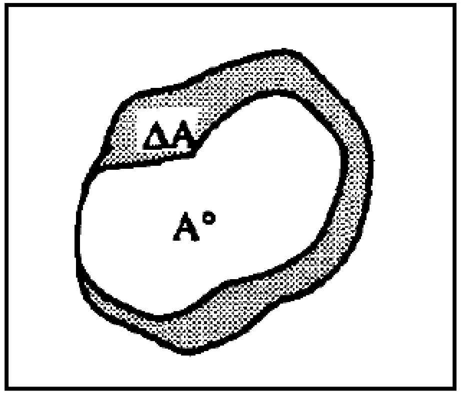


Matrices avec les frontières certaines :

$$M = \begin{pmatrix} A^\circ \cap B^\circ & A^\circ \cap \partial B & A^\circ \cap B^- \\ \partial A \cap B^\circ & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^\circ & A^- \cap \partial B & A^- \cap B^- \end{pmatrix}$$

8 relations entre 2 régions simples certaines

(Clementini et al., IJAR, 1997)



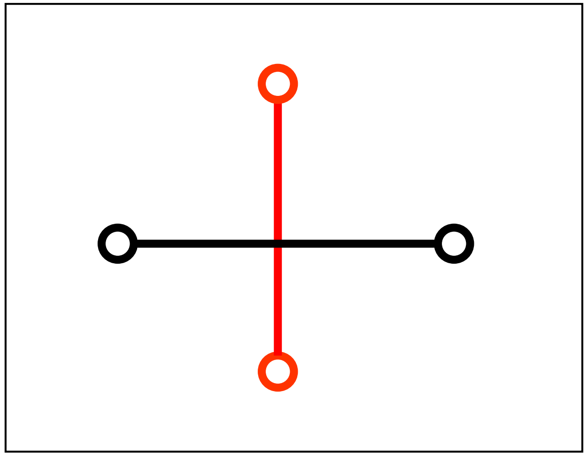
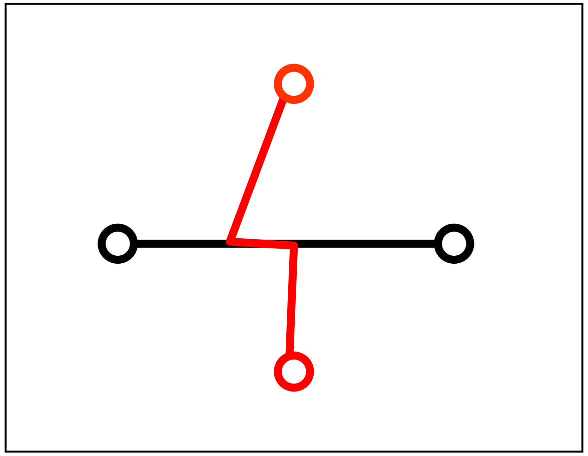
Matrices avec les frontières larges :

$$M = \begin{pmatrix} A^\circ \cap B^\circ & A^\circ \cap \Delta B & A^\circ \cap B^- \\ \Delta A \cap B^\circ & \Delta A \cap \Delta B & \Delta A \cap B^- \\ A^- \cap B^\circ & A^- \cap \Delta B & A^- \cap B^- \end{pmatrix}$$

44 relations entre 2 régions simples incertaines

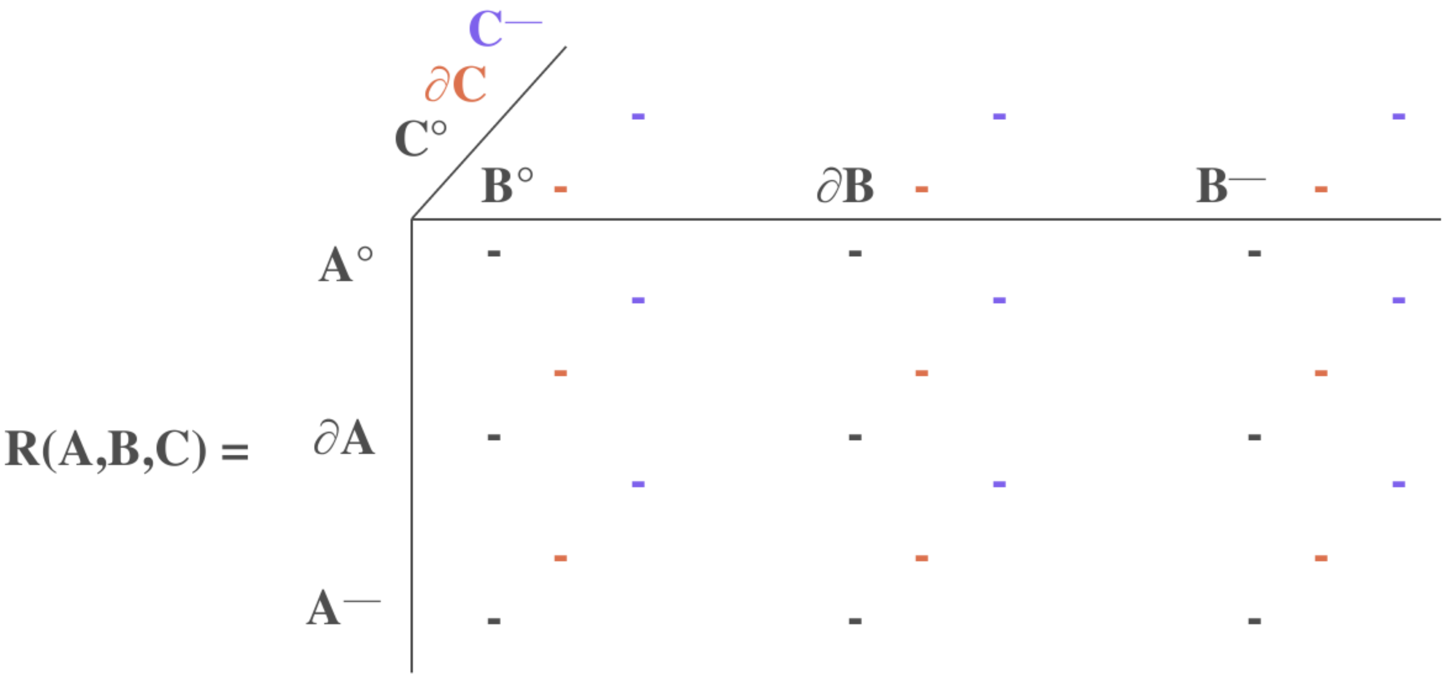
Vers d'autres modèles de relations topologiques

Dimensionally Extended-9IM permet de faire de nouvelles nuances :

		
9IM	$\left[A^\circ \cap B^\circ \neq \varnothing \right] = 1$	$\left[A^\circ \cap B^\circ \neq \varnothing \right] = 1$
DE-9IM	0D	1D

Vers d’autres modèles de relations topologiques

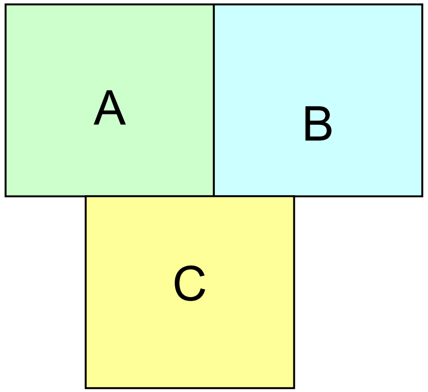
relation ternaire = relation entre 3 objets



(Favetta, Thèse, 2003)

Vers d'autres modèles de relations topologiques

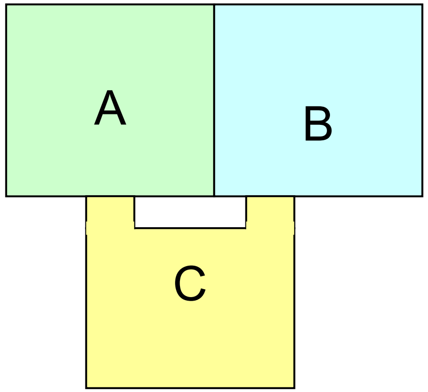
Si on ne considère que des relations binaires :



$$\begin{bmatrix} \delta A \cap \delta B \neq \varnothing \end{bmatrix} = 1$$

$$\begin{bmatrix} \delta A \cap \delta C \neq \varnothing \end{bmatrix} = 1$$

$$\begin{bmatrix} \delta B \cap \delta C \neq \varnothing \end{bmatrix} = 1$$



$$\begin{bmatrix} \delta A \cap \delta B \neq \varnothing \end{bmatrix} = 1$$

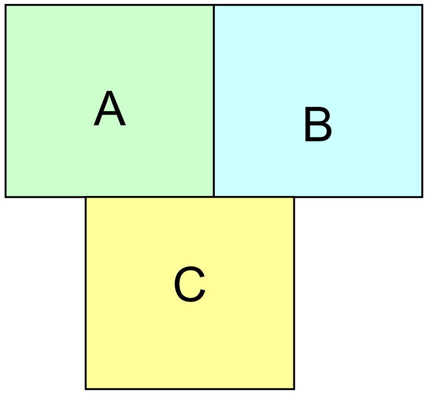
$$\begin{bmatrix} \delta A \cap \delta C \neq \varnothing \end{bmatrix} = 1$$

$$\begin{bmatrix} \delta B \cap \delta C \neq \varnothing \end{bmatrix} = 1$$

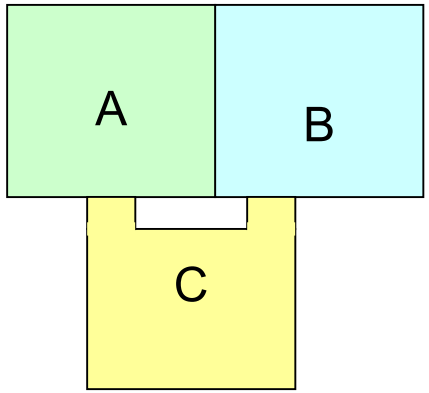
Vers d'autres modèles de relations topologiques

relation ternaire

Si on considère des relations ternaires :



$$\left[\delta A \cap \delta B \cap \delta C \neq \emptyset\right] = 1$$



$$\left[\delta A \cap \delta B \cap \delta C \neq \emptyset\right] = 0$$

Test de la cohérence des relations

On définit des contraintes d'intégrité.

Mais elles peuvent être incohérentes

Exemple d'un ensemble incohérent de relations

Conclusion

Le 13 novembre à 14h à Irstea

Un atelier conjoint sur la qualité des données, entre le GT Big data et l'action « incertitude » du GDR MAGIS.