Geographic Knowledge for Territorial Intelligence

Pr. R. Laurini
INSA, University of Lyon
KSI Fellow, Chicago

Geographic Knowledge for Territorial Intelligence

- 1 – Territorial Intelligence
- 2 – Generalities about Geographic Knowledge
- 3 – Generic Geographic knowledge
  - 3.1 – Mutation of topological relations
  - 3.2 – Gazetteers and toponyms/placenames
- 4 – Modeling principles
- 5 – Conclusions

Geographic relations

- In addition to spatial relations
  - Tessellations for administrative objects
  - Networks
  - Ribbon relations
  - Geographic ontologies with Geo Relations
  - Gazetteers

1 – Territorial Intelligence

- Business intelligence applied to territories
  - Cities (→ smart cities)
  - Regions, Countries
- Links with urban, regional and environmental
  - Planning
  - Management
- Objective: Sustainable development

A new family of concepts

- Such as
  - competitive intelligence,
  - strategic economic intelligence,
  - distributed intelligence,
  - social intelligence, or collective,
- emphasizing organized and systematic collection, analysis and dissemination of information for the purpose of development.

Territorial Intelligence

Territorial Intelligence

=

(Territory

+

Collective Human Intelligence

+

Artificial Intelligence)

➔ Sustainable development)

2 – Generalities about GK

- Definitions
  - Feature = geographic entity existing in the real word
  - Geographic object = computer representation of a feature
  - Rule = mathematical inference
- Not only logics, but also space/geometry

AI + Computational Geometry

- Necessity to include
  - Computational geometry
  - Topology
  - Spatial analysis
  - Operation research
  - Linguistics
  - Etc.
- Earth rotundity

Generic and specific knowledge

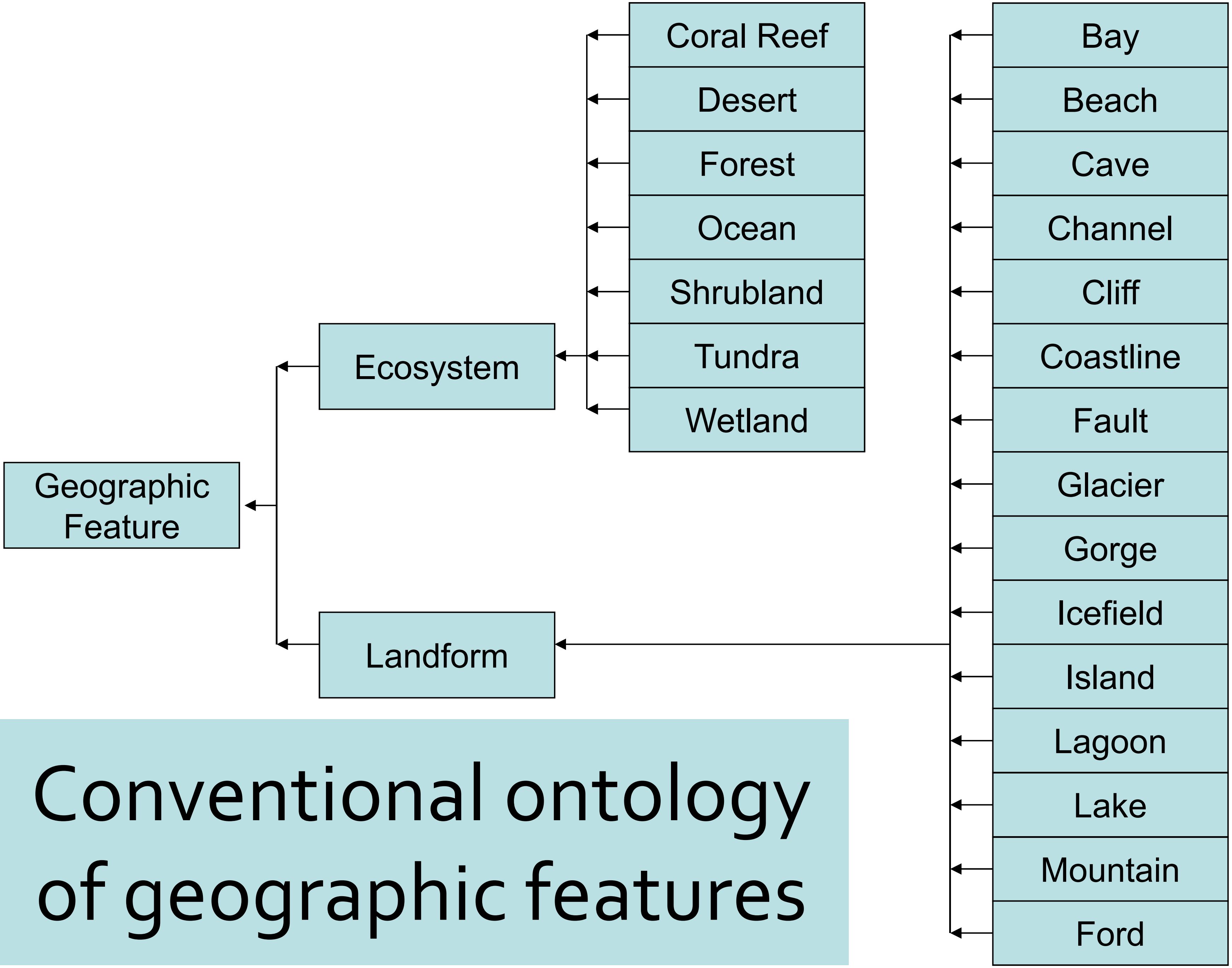
- Specific knowledge
  - Devoted to a particular place in the world
  - F.i. Antarctica, near Equator, etc.
  - Mountains, seashore
- Generic knowledge
  - Valid everywhere
  - Links with acquisition devices
  - Links with maths and linguistics

Application knowledge

- Knowledge rules valid in one domain
  - Urban planning
  - Environmental planning
  - Transportation, logistics
  - Etc.

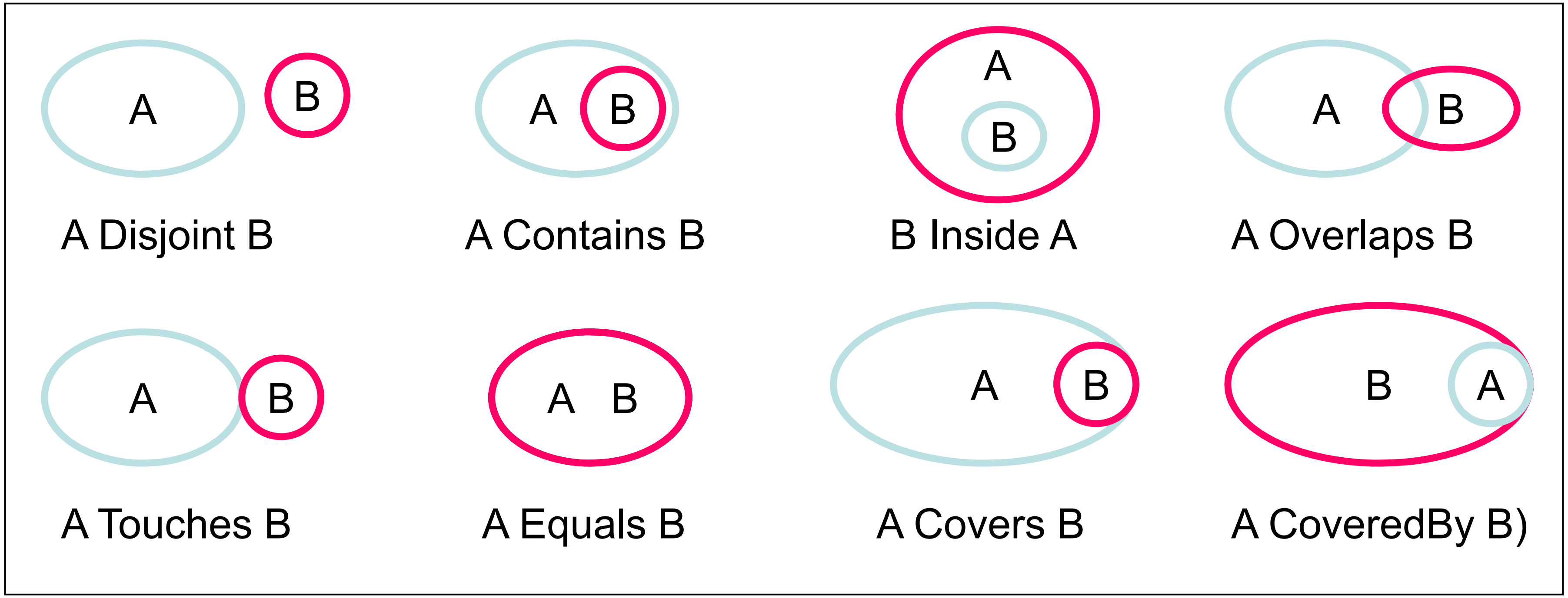
Geographic Ontologies

- Organizations of geo features
- Relations « *is_a* », « *has_a* », « *whole_part* »
- Necessity of spatial relations

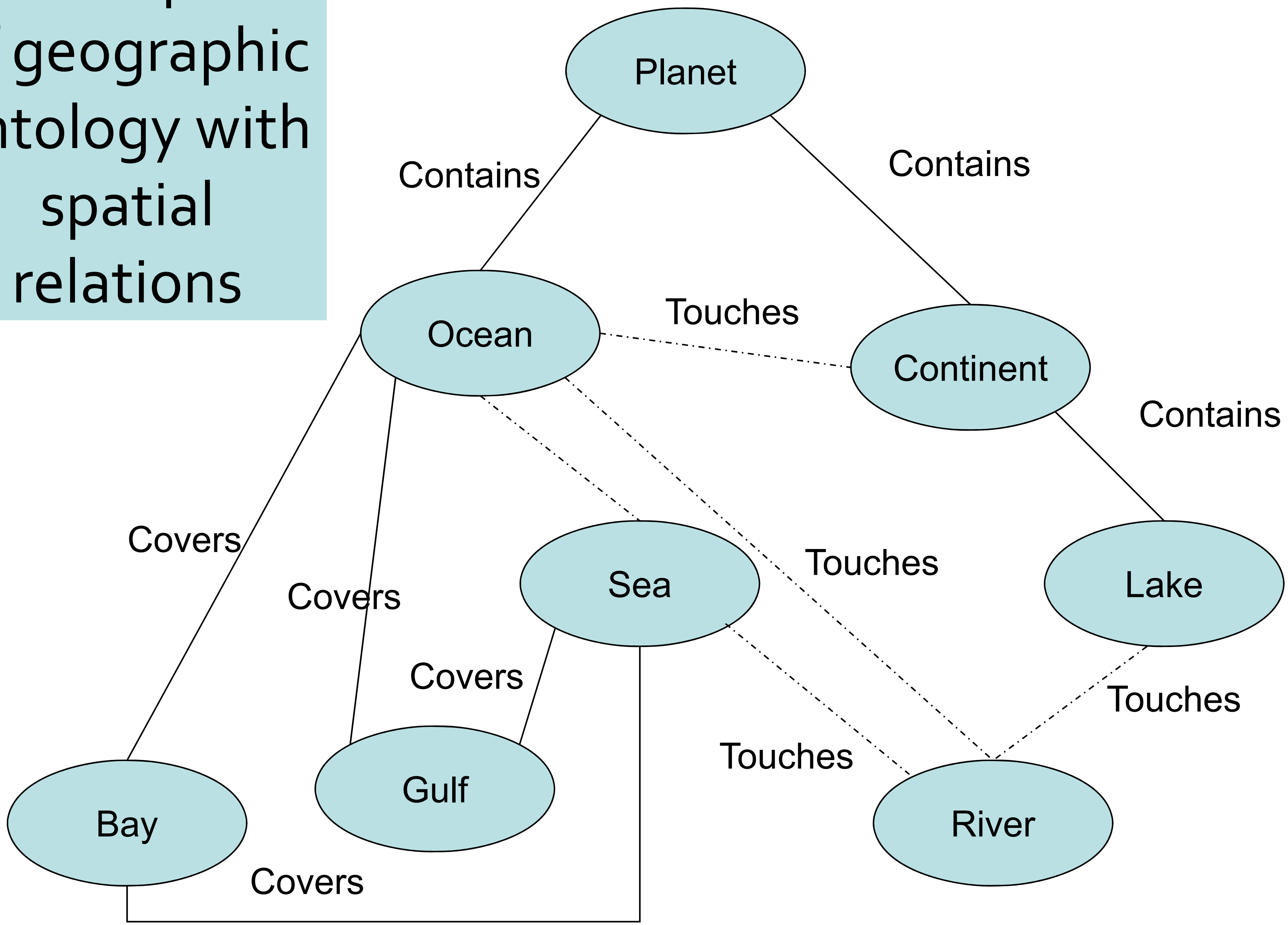


Conventional ontology of geographic features

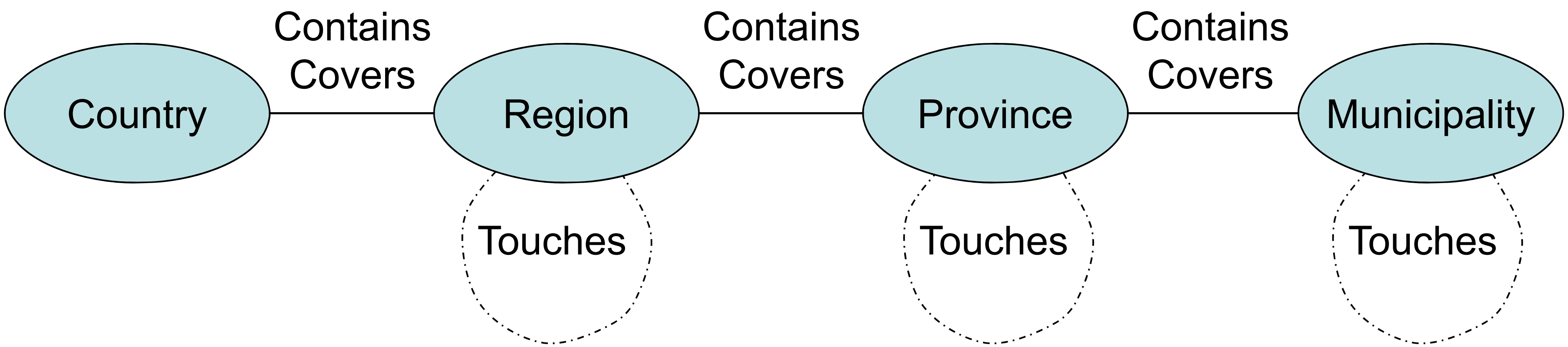
Egenhofer topological relations



Example of geographic ontology with spatial relations



Administrative subdivisions



Rapid analysis of toponyms

- “Mississippi” can be the name of a river or of a state.
- The city, “Venice”, Italy, is also known as “Venezia”, “Venise”, “Venedig”, respectively, in Italian, French and German.
- The local name of the Greek city of “Athens” is “Αθήνα”; read [a’θina].
- “Istanbul” was known as “Byzantium” and “Constantinople” in the past.
- The modern city of Rome is much bigger than in Romulus’s time.
- There are two Georgias, one in the United States and another one in Caucasia.

- The toponym “Milano” can correspond to the city of Milano or the province of Milano.
- The river “Danube” crosses several European countries; practically in each country, it has a different name, “Donau” in Germany and Austria, “Dunaj” in Slovakia, “Duna” in Hungary, “Dunav” in Croatia and Serbia, “Dunav” and “Дунав” in Bulgaria, “Dunărea” in Romania and in Moldova and “Dunaj” and Дунай” in the Ukraine. It is also called “Danubio” in Italian and Spanish, “Tonava” in Finnish and “Δούναβης” in Greek.

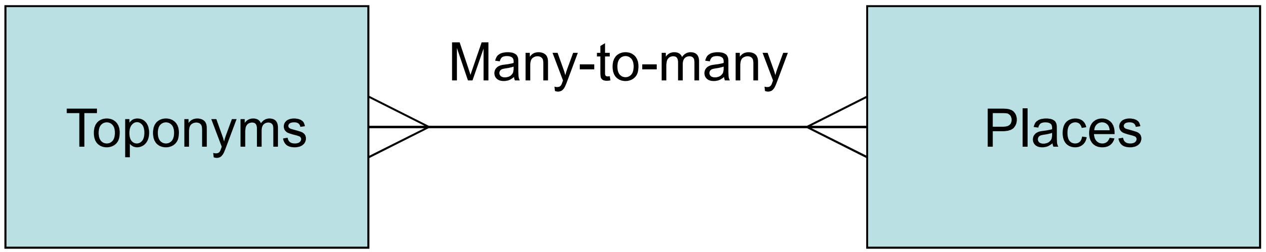
- Sometimes, names of places can be also names of something else; for instance “Washington” can also refer to George Washington or anybody with this first name or last name.
- In the U.K., there are several rivers named Avon.
- Some place names are formed of two or several words; for instance, “New Orleans”, “Los Angeles”, “Antigua and Barbuda”, “Trinidad and Tobago”, “Great Britain”, “Northern Ireland.

- Some very long names can have simplifications; the well-known Welsh town “Llanfairpwllgwyngyllgogerychwyrndrobwllllantysiliogogogoch” is often simplified to “Llanfair PG” or “Llanfairpwll”.
- Some abbreviations can be common, such as “L.A.” for “Los Angeles”, whereas its name at its inception was “El Pueblo de Nuestra Señora la Reina de los Angeles del Rio de la Porciúncula”;

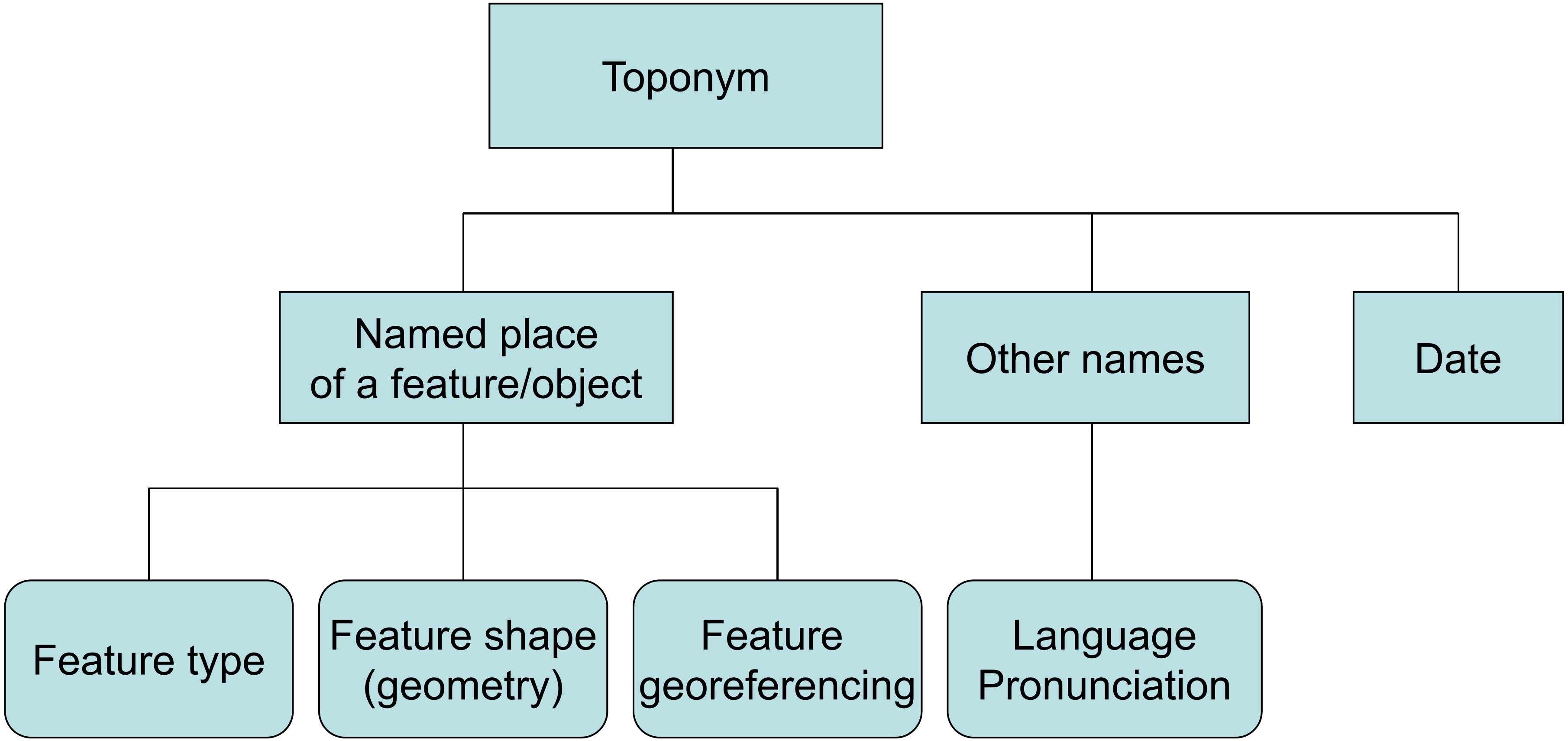
- Peking became Beijing after a change of transcription to the Roman alphabet; but the capital of China has not modified its name in Chinese.
- Declensions: (der Rhein, des Rheins, etc.).

Gazetteers

- A dictionary of toponyms/placenames
- A database structure for placenames

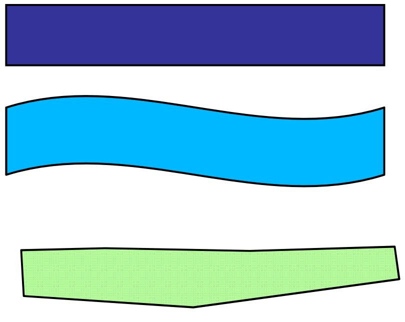


Example of gazetteer

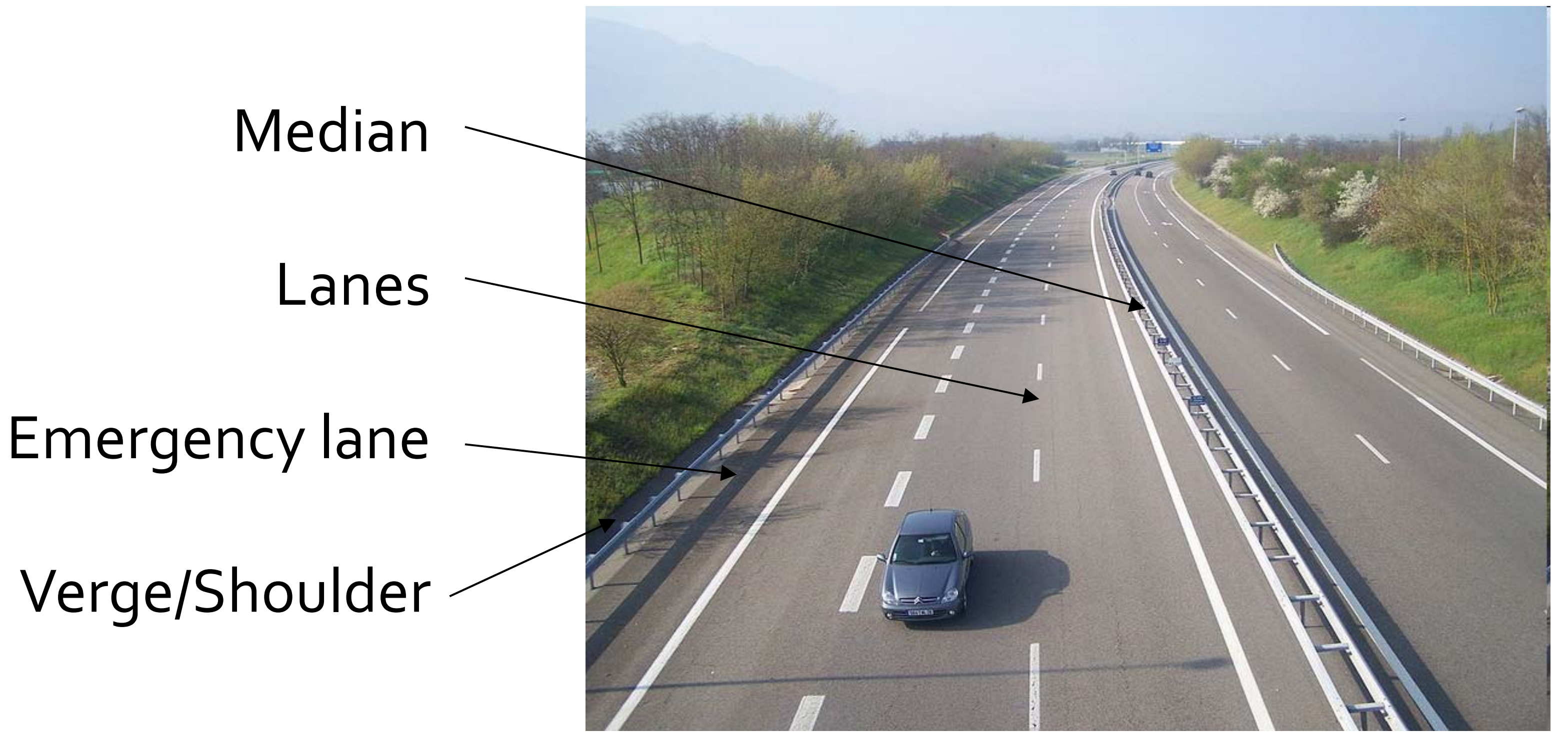


About ribbons

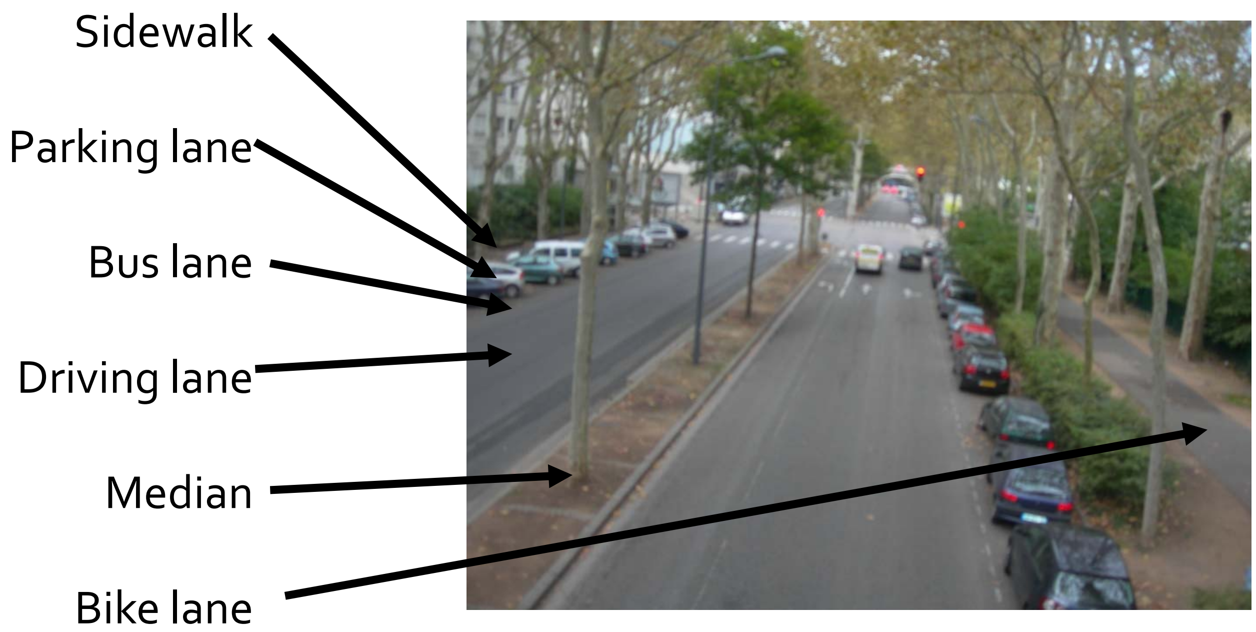
- Line with some width
- ➔ area
- Rectangular ribbon
- Extended ribbon
- Loose ribbon
- Relations between ribbons



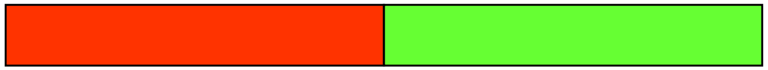
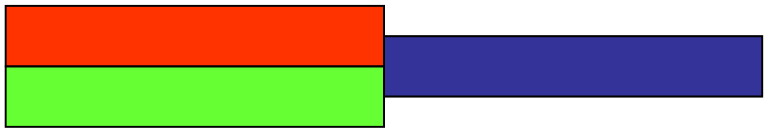
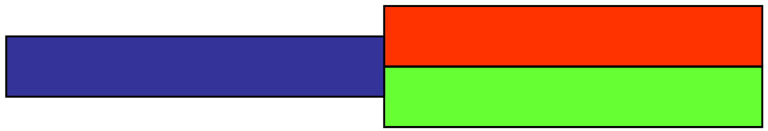
Modeling with ribbons



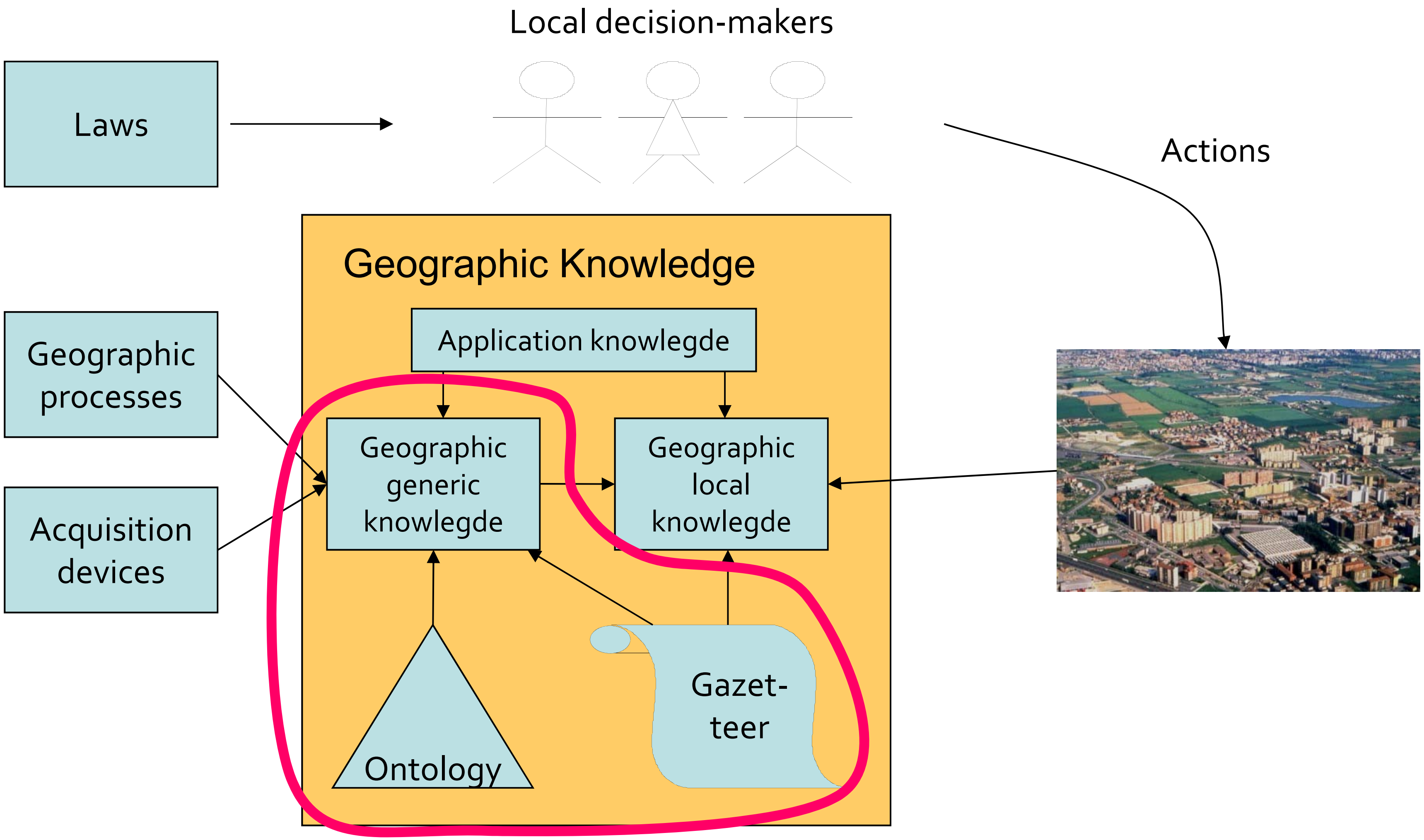
Urban example



Relations between ribbons

Side-by-side	
End-to-end	
Fusion	
Splitting	

Organization of Geographic Knowledge

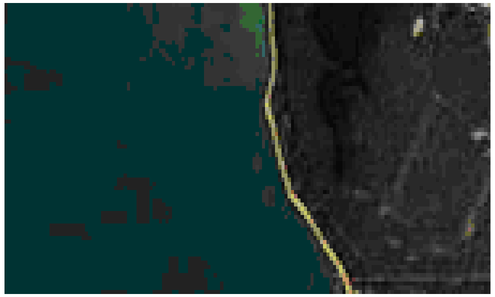


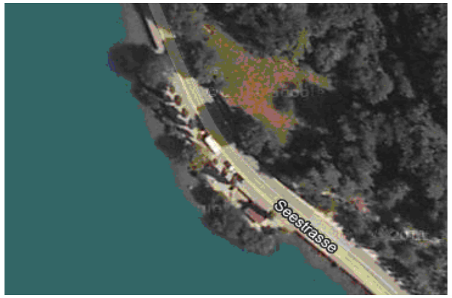
3 – Generic Geographic knowledge

- Geographic knowledge valid everywhere
- Linked to
  - Maths
  - Linguistics
  - Acquisition devices
- Only three types
  - Mutation of topological relations
  - Gazetteers and toponyms
  - About raster reasoning

3.1 – Mutation of topological relations

- Granularity of interest
- Independence from scale
- Ex. Road along a coast
  - Touches
  - Disjoint
- According to scales, topological relations can vary





Visual acuity

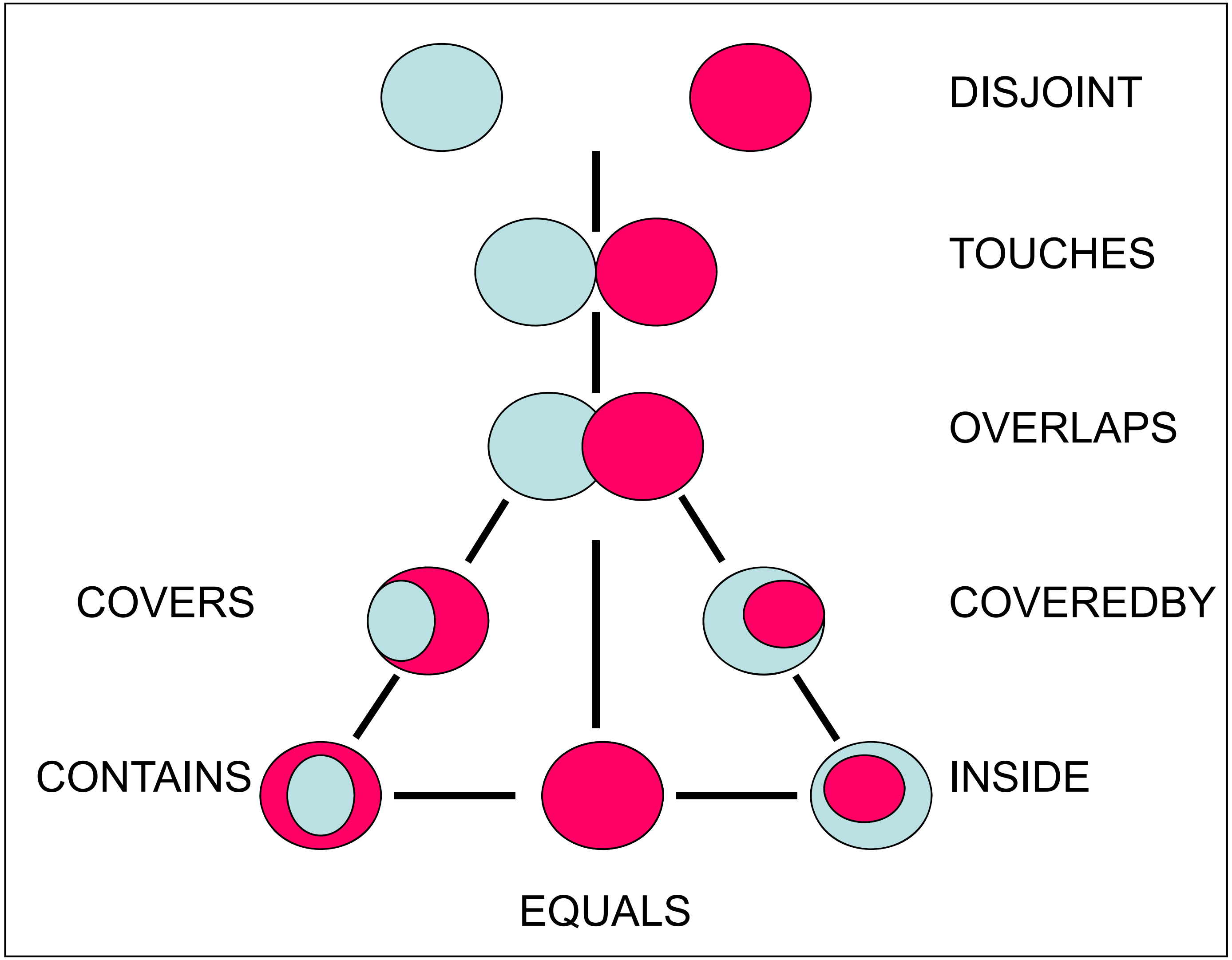
- According to scale, objects are present or not.
- Cities: area, then point, then nothing
- River: ribbon, then line, then nothing
- Threshold for visual acuity
  - 0.1 mm (object no more visible)
  - 1 mm (ribbon is transformed into a line)

Features and mutation of their corresponding geographic objects

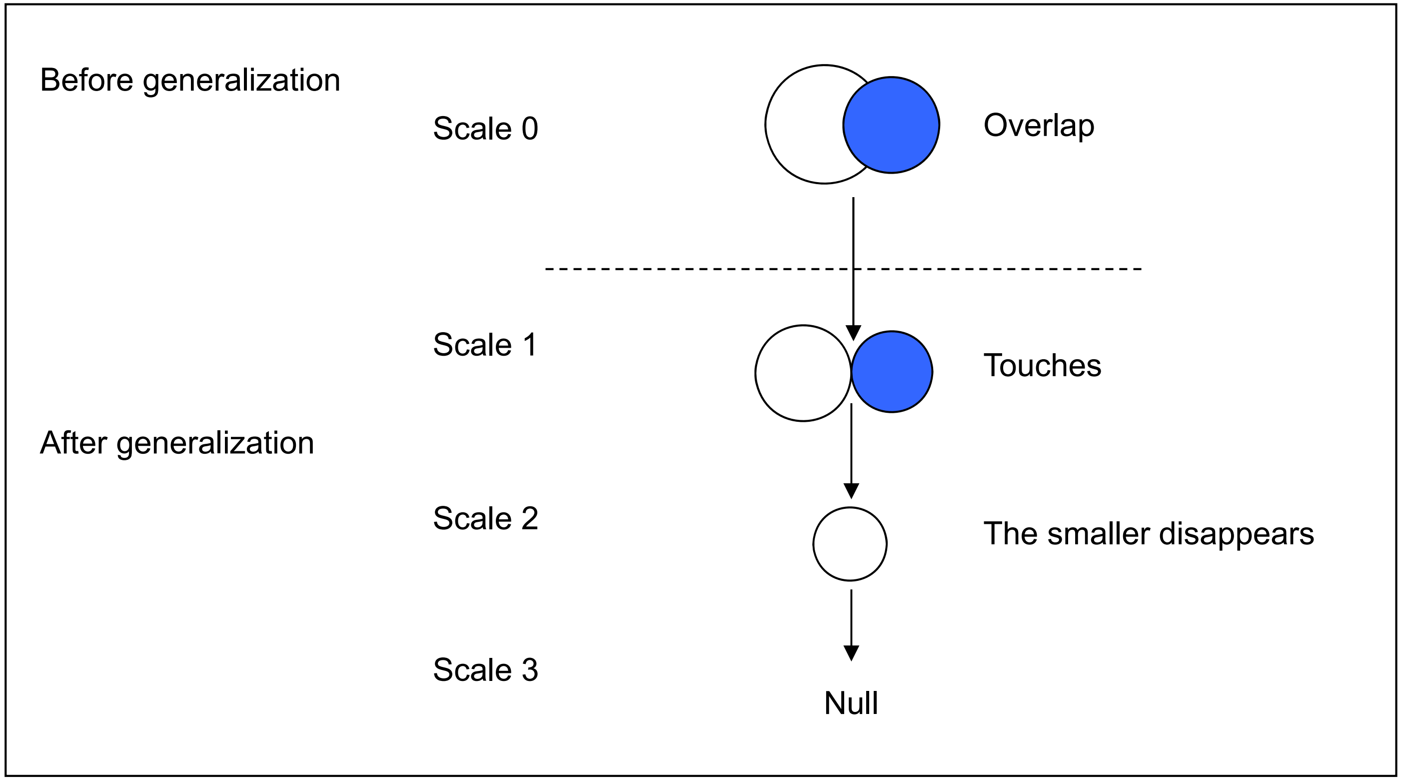
100 Km wide city	Invisible	Reduced to point	Visible area
1 ha wide hamlet	Invisible	Reduced to point	Visible area
100 m wide motorway	Invisible	Reduced to line	Visible ribbon
1 m wide path	Invisible	Reduced to line	Visible ribbon

10^{-10} 10^{-8} 10^{-6} 10^{-4} 10^{-2} 1 Logarithmic Scale

Vicinity of topological relations



Example: Overlap-to-Touches



Ex. From OVERLAPS to TOUCHES

$$\forall \mathrm{O}^1, \mathrm{O}^2 \in \mathrm{GeObject}, (\forall \sigma \in \mathrm{Scale})$$
$$\wedge (O_\sigma^1 = 2Dmap(O^1, \sigma)) \wedge (O_\sigma^2 = 2Dmap(O^2, \sigma))$$
$$\wedge (Overlaps(\mathrm{O}^1, \mathrm{O}^2)) \wedge (Area(\mathrm{O}^1 \bigcap \mathrm{O}^2) < Area(\neg(\mathrm{O}^1 \bigcap \mathrm{O}^2)))$$
$$\Rightarrow Touches(O_\sigma^1, O_\sigma^2).$$

In which *2Dmap* is a cartographic function

Other possible mutations

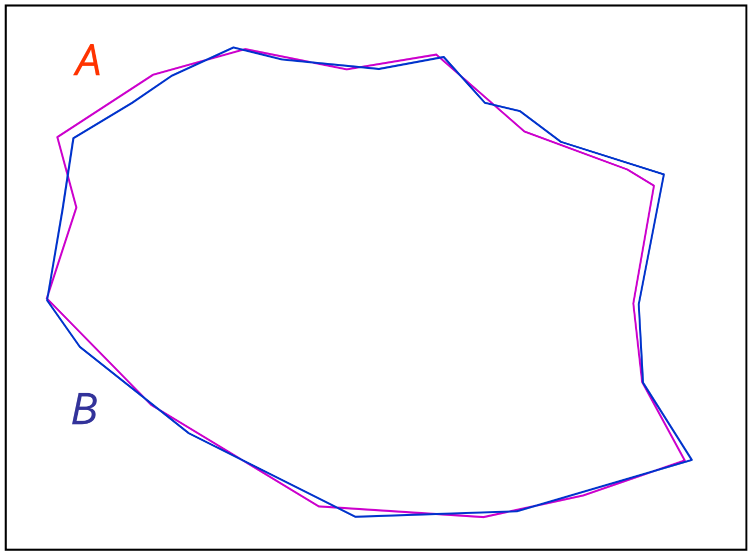
- Disjoint-to-Touches
- Overlaps-to-Covers
- Contains-to-Touches

3.2 – Gazetteers and toponyms

- Geographic information retrieval
- Multilingualism
- Concepts in different languages are different

Geometric homology

- Taking measurement errors or uncertainties into account
- The same feature can have different geometric representations
- Geometry: polygons with point coordinates



Linguistic homology

Homology class
(Objects linked by ₪)

Equivalence class
(Objects linked by ≡)

- United States of America
- United States
- USA
- U.S.A.

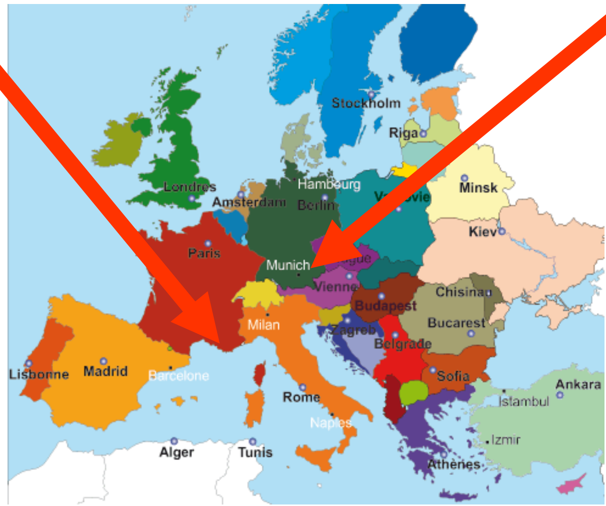
- Etats Unis d'Amérique
- Estados Unidos de América
- Stati Uniti d'America
- Vereinigte Staaten von Amerika
- Соединённые Штаты Америки
- etc

Non-linguistic transitivity

Principauté de Monaco
Múnegu
Μονακό
Монако

Monaco

Munich
München
Minga
Μόναχο
Мюнхен
Monaco di Baviera



Type homology

- Consider two geographic ontologies in different languages
- Equivalence or homology
- Example: French « quai », three meanings

  - Wharf ₪*muelle*
  - Riverside ₪*avenida a lo largo de un río*
  - Platform ₪*andén*

Formalization

- Set of languages: $\lambda \in \Lambda$
- Ontology of types: Ω = set of *Types* with relations between them
- Gazetteer: Γ = set of *Toponyms*
- Set of spatial relations
- Geometric Earth: *Geoid*

Definition of GK System

- $GKS = \{T, \lambda, \Omega, \Gamma, Og, \mathscr{R}\}$
  - T Inside Geoid
  - $\lambda \in \Lambda$
  - $Og=\{Og^1,\ldots Og^n: n \in N\}$
- $Og^i = (id^i, geom^i, Type^i, Toponym^i]$
  - $Type^i \in \Omega$
  - $Toponym^i \in \Gamma$
- $\mathscr{R}$ set of relationships $\{Og^i \; R \; Og^j: (i, j \leq n) \wedge (i, j \in N)\}$
- R relation

Considering 2 GKS

- $GKS_1 = \{T_1, \lambda_1, \Omega_1, \Gamma_1, Og_1, \mathscr{R}_1\}$
- $GKS_2 = \{T_2, \lambda_2, \Omega_2, \Gamma_2, Og_2, \mathscr{R}_2\}$
  - *With*
  - $T_1 \cap T_2 \neq \emptyset$
  - $\lambda_1 \neq \lambda_2$
  - $\Omega_1 \neq \Omega_2$
  - $\Gamma_1 \neq \Gamma_2$

Same features
But different geographic objets
And maybe different relationships

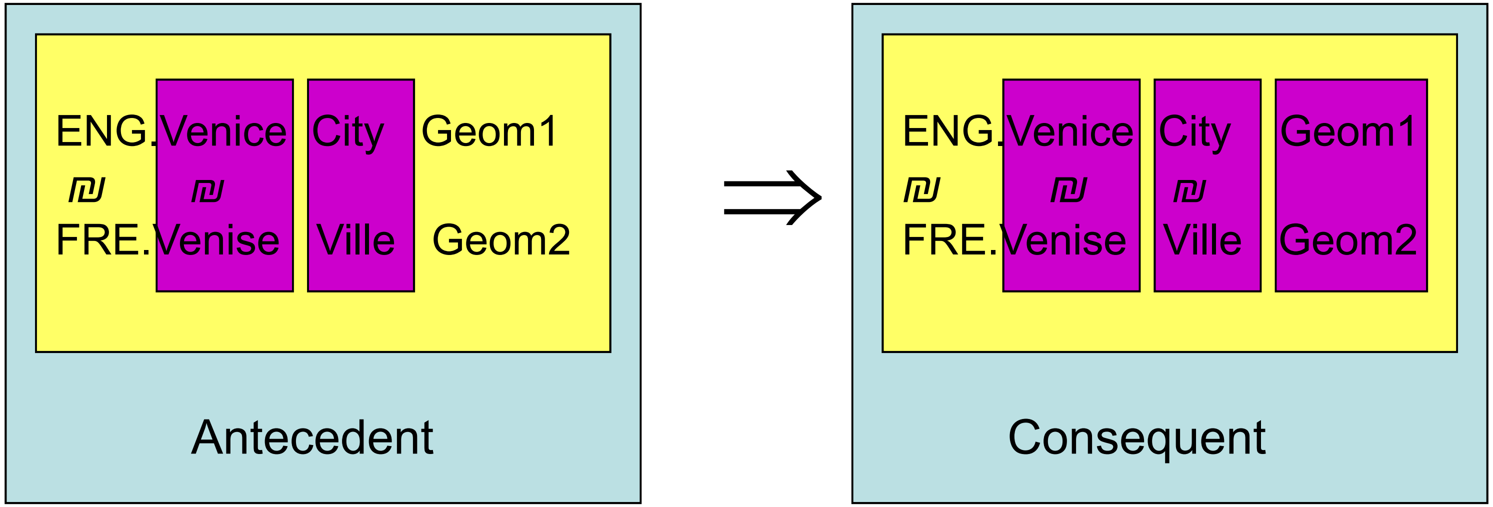
Inferring geometry: Rule #1

IF

Homologous toponyms AND
Homologous types

THEN

Homologous geometry



Rule 1bis: with MBR

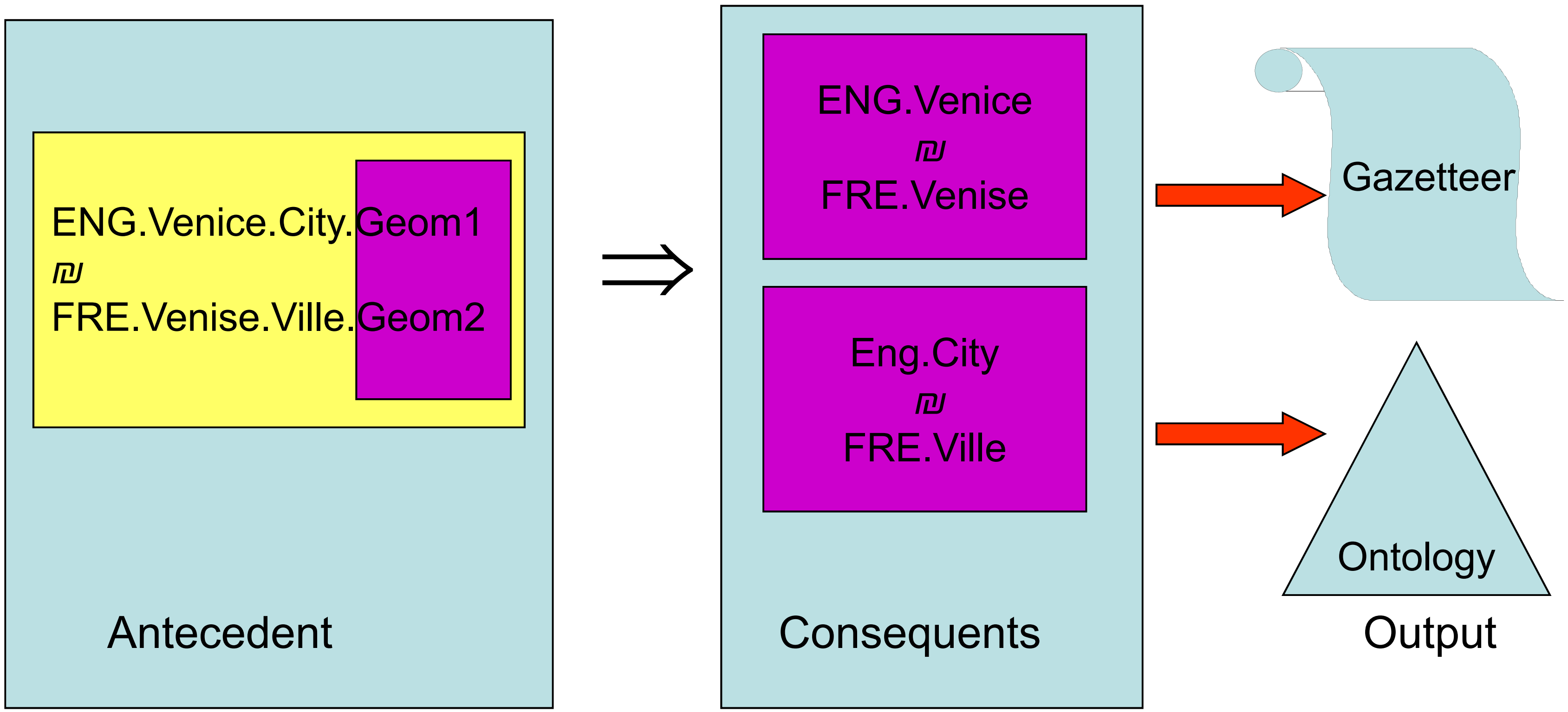
Inferring types and toponyms Rule #2

IF

Homologous geometry

THEN

Homologous toponyms
homologous types



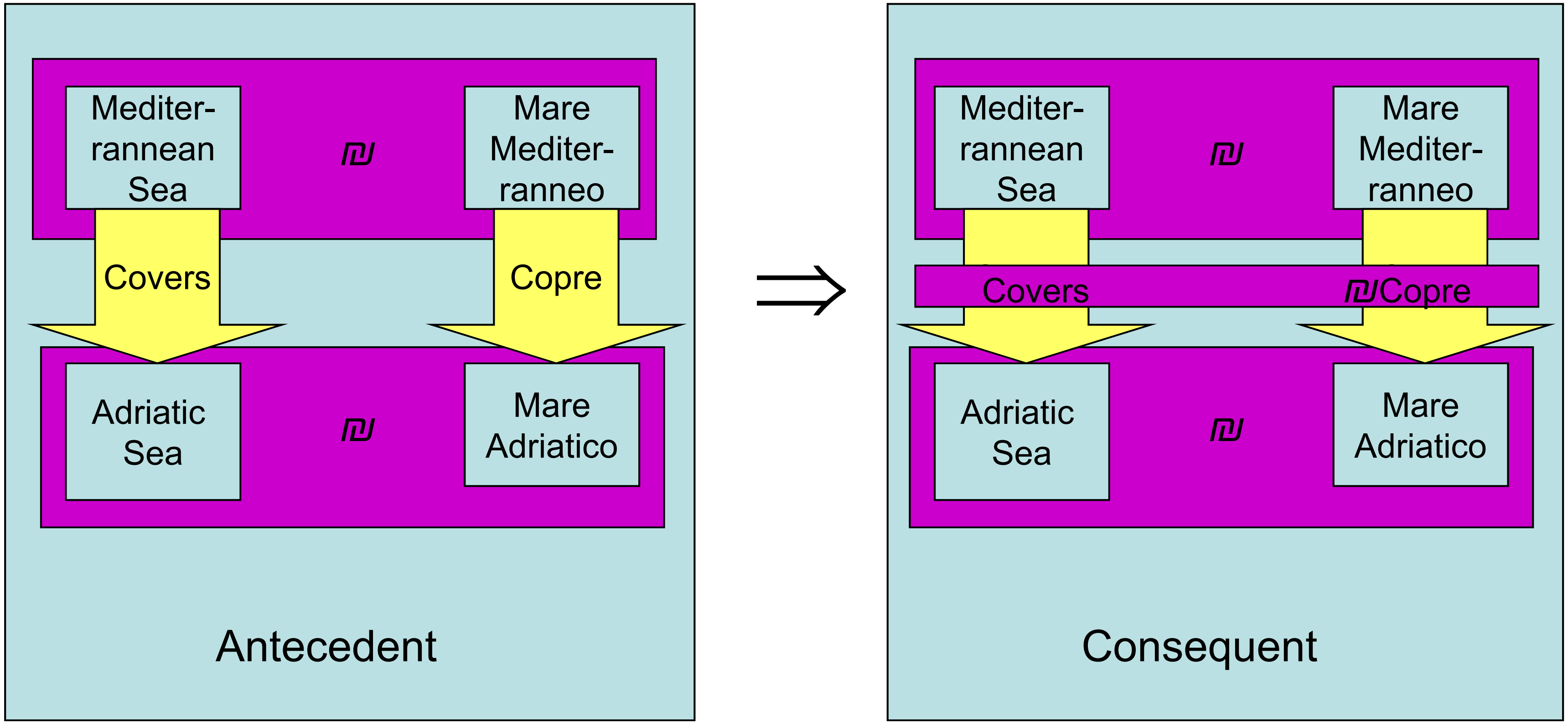
Rule #3

IF

Two pair of homologous geographic objets are linked thru 2 different relations

THEN

Those relations are homologous



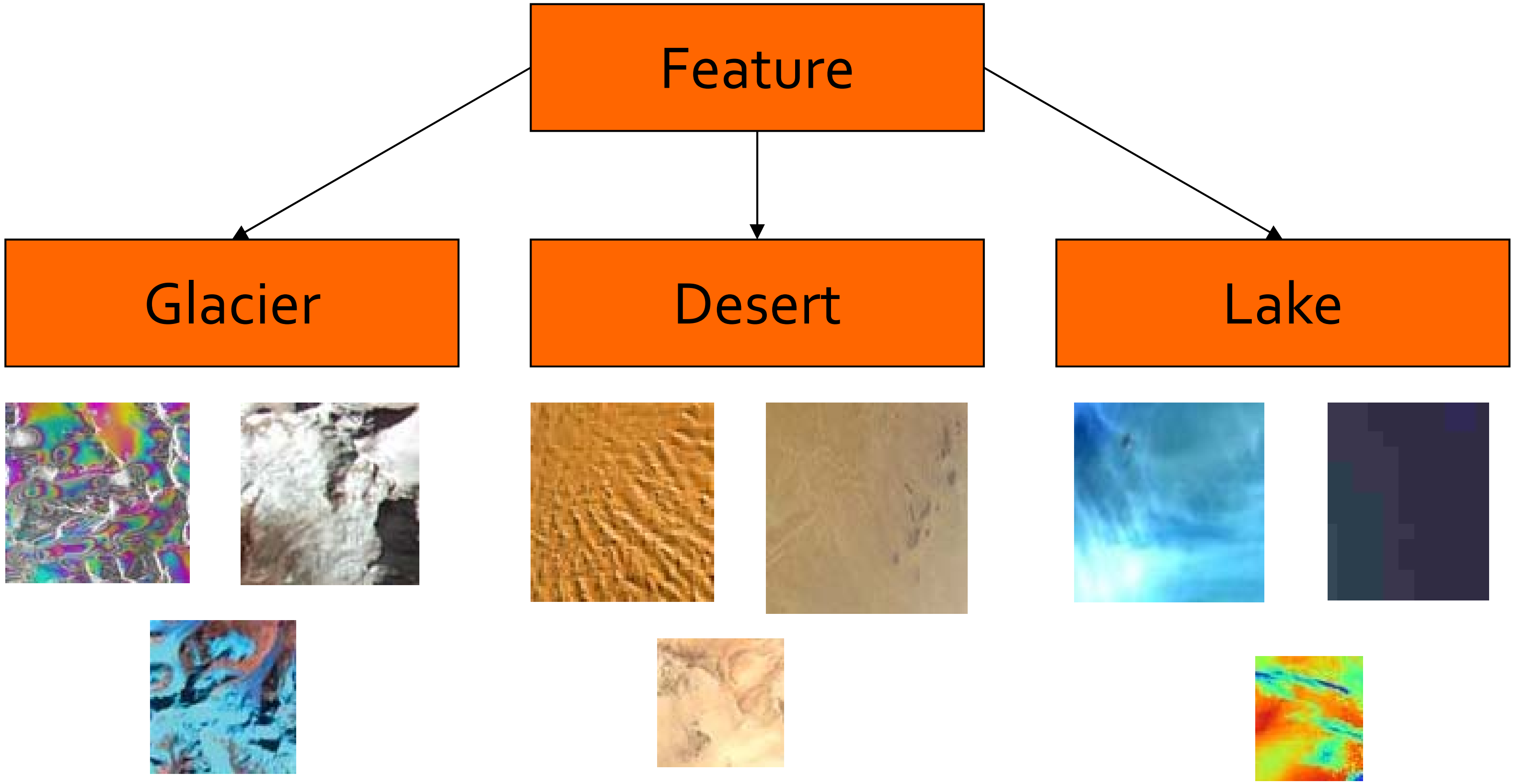
3.3 – From raster representation

- Aerial photos / Satellite images
- Analysis
  - Pattern recognition
- Usage
  - Feature recognition
  - Updating

Visual Ontology

- In addition to feature type, several samples taken at different scales
- Samples = pictures from sky
  - Hyperspectral
  - Different channels
  - Different focus
  - Etc.
- Pictures only at lower level

Excerpt of a visual ontology



3.4 General characteristics

- Geographic knowledge reasoning
  - Independence from scale
  - Independence from data acquisition techniques
  - Independence from languages
  - Easy integration of
    - Spatial analysis
    - Network analysis

4 – Modeling principles

- Theoretical bases for modeling geographic knowledge
- 12 principles and 12 prolegomena
- Prolegomena: preliminary considerations

Construction of this framework

- More than 30 years of teaching GIS
- Necessity to reorder GIS concepts
- Necessity of testing this framework
  - Expert consensus
- First presentations
  - Belluno (2/2012), Salerno (3/2012), Sousse (6/2012), Dublin (2/2013),
- Brighton (7, 2013)
  - Kuala Lumpur (9, 2013)

Prolegomenon #1
(3D +T objects)

- ***"All existing objects are tridimensional and can have temporal evolution; lower dimensions (0D, 1D and 2D) are only used for modeling (in databases) and visualization (cartography)".***

- Unlike geodetic objects which were created by man, all features are 3D, can move, can change their shape and can be destroyed.

Prolegomenon #2 (acquisition by measurements)

- ***"All basic attributes (spatial or non-spatial) are obtained by means of measuring apparatuses having some limited accuracy".***
- Now more and more data come from sensors;
- more, citizens can be seen as sensors

Prolegomenon #3
(Continuous fields)

- "***Since it is not possible to store the infinite number of value points in a continuous field, some sampling points will used to generate the whole field by interpolation.***

Prolegomenon #4 (Raster-vector and vector-raster transformations)

- ***"Procedures transforming vector-to-raster data and raster-to-vector data must be implemented with loosing less accuracy as possible".***

Prolegomenon #5

(From Popper's falsifiability principle):

- ***"When a new apparatus delivers measures with higher accuracy, these measures supersede the previous ones".***
- The practical consequence is that as a new generation of data comes, geographic data and knowledge basis must integrate those data. But alas, due to the acquisition cost, a lot of actual systems are based on "obsolete" data.

Prolegomenon #6 (Permanent updating)

- ***"Since objects are evolving either continuously (sea, continental drift) or event-based (removing building), updating should be done permanently respectively in real-time and as soon as possible".***

- Sensor-based updating
- Data cleaning / Data quality

Prolegomenon #7 (Geographic metadata)

- ***"All geographic databases or repositories must be accompanied with metadata".***
- International Standard ISO 19115 "Geographic Information - Metadata" from ISO/TC 211 provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.
- Practically, many geographic databases do not implement the whole standard, but only the more important aspects, because it is very time-consuming.

Prolegomenon #8
(Cartographic objects)

- ***"In cartography, it is common to eliminate objects, to displace or to simplify them".***

- This is due to ensure a maximal readability of maps.

Prolegomenon #9
(One storing, several visualizations)

- ***"A good practice should be to store all geographic objects with the highest possible accuracy and to generate other shapes by means of generalization".***

- This can be seen also as a consequence of Prolegomenon #3.

Prolegomenon #10 (Place names and gazetteers)

- ***"Relationships between places and place names are many-to-many".***

- Mississippi is the name of a river and the name of a state. The actual city of Rome, Italy, is larger than the same Rome in Romulus's time.

- The main consequence is that unique feature identifiers are not so easy to define.

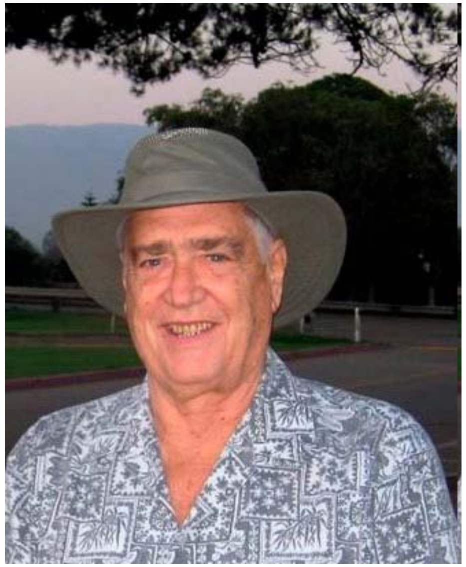
Prolegomenon #11 (Geographic ontologies)

- ***"All geographic object types are linked to concepts organized into a geographic ontology based on topological relations".***

Prolegomenon #12 (Tobler's law):

- ***"Everything is related to everything else, but near things are more related than distant things".***

This statement may be seen as a key-concept also for geographic data mining



Principles

- Prolegomena: preliminary considerations
- Principles
  - Basis for modeling geographic knowledge
  - Basis for transforming it

Principle #1
(Origin of geographic knowledge):

- ***"Spatial knowledge is hidden in geometry whereas geographic knowledge comes in addition from non-spatial attributes".***

Principle #2 (Knowledge cleaning)

- ***"All geographic data, once captured, must be cleaned to remove errors and artifacts".***
- All automatic acquisition system may include errors or anomalies.
- Be aware when generating knowledge!

Principle #3
(Knowledge enumeration)

- ***"It is not necessary to enumerate all possible chunks of geographic knowledge".***
- if one has *n* object, then $(n\text{-}1)^2$ North-South relationships can be also derived accordingly.

Principle # 4
(From geoid to plane):

- ***"On small territories, a planar representation is sufficient whereas for big territories, Earth rotundity must be taken into consideration".***
- But the question is "how to define a small or a big territory"?
- A solution can be to define a threshold, for instance a 100 km wide square.

Principle #5
(Visualization and visual acuity)

- ***"Cartographic representation is linked to visual acuity".***

- Thresholds must be defined. In classical cartography, the limit ranges from 1 mm to 0.1 mm.

Modification

- Disappearance
  - $\forall O \in GeObject, \forall \sigma \in Scale$
  - Ù O_σ =2Dmap(O, σ)
  - Ù Area $(O_\sigma)< \varepsilon_2$
  - => $O_\sigma = \emptyset$.
- Transformation into point
  - $\forall O \in GeObject, \forall \sigma \in Scale$
  - Ù O_σ =2Dmap(O, σ)
  - Ù ε_1 Area $(O_\sigma) < \varepsilon_2$
  - => O_σ = Centroid(O).

Principle #6 (Sharpification)

- ***"At some scales* every *fuzzy object becomes sharp*".**
- Egg-yolk representation
  - When the mean distance between egg and yolk is less than a threshold
  - Its geometry can be taken midway

Principle #7
(Relativity of spatial relations)

- **"*Spatial relation varies according to scale*".**

- $\forall O^1, O^2 \in GeObject, \forall \sigma \in Scale$
- Ù O^1_σ =2Dmap(O^1, σ) Ù O^2_σ =2Dmap(O^2, σ)
- Ù Disjunct (O^1, O^2)
- Ù Distance $(O^1, O^2) < \varepsilon_1$
- => Touch (O^1_σ, O^2_σ)

Principle #8
(Transformation into graphs)

- "***Every set of linear objects can be transformed into a graph***".

- For instance from
  - Roads to road networks
  - Rivers to river graphs

Principle #9
(From pictorial to geographic objects)

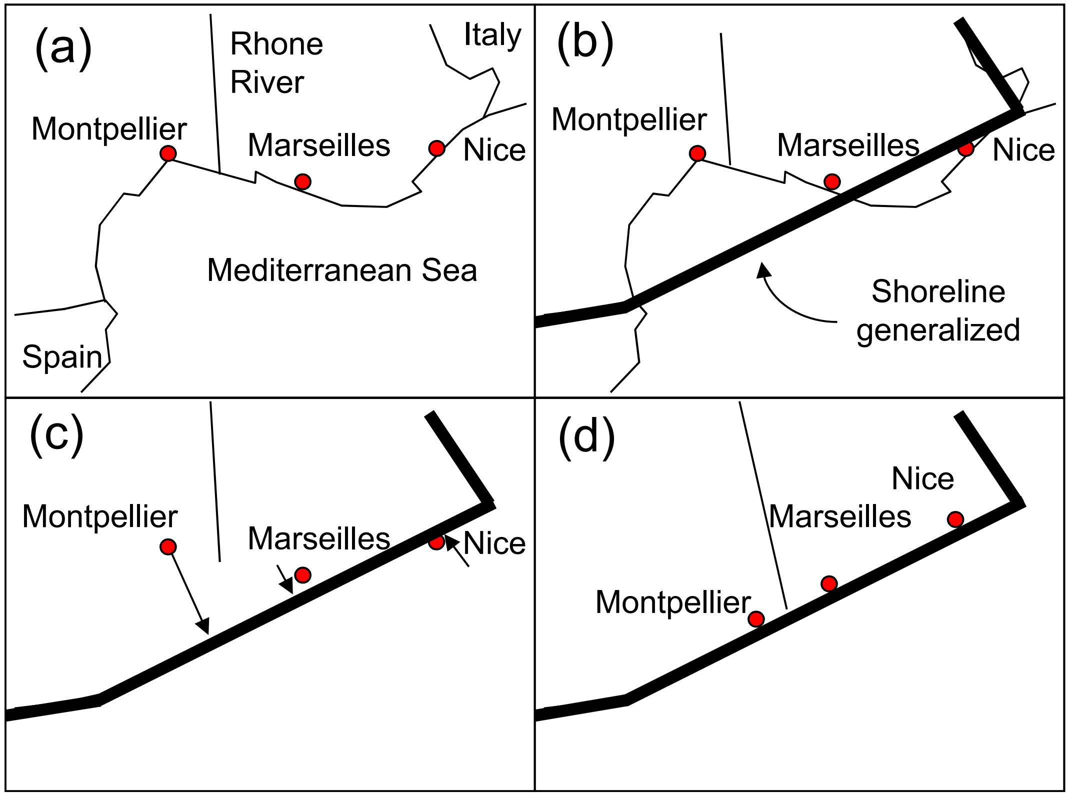
- ***"Any group of pixels having same characteristics can be regrouped into a pictorial object; this pictorial object can be conferred a geographic type possibly using an ontology".***
- Indeed as soon as a pictorial object is recognized its type will be identified and it can be a part of a geographic object.

Principle #10
(Visualization constraints)

- ***"The spatial relations between objects must hold after generalization".***

- Ex. Mediterranean coastline

Example


(a)
Rhone
River
Italy
Montpellier
Marseilles
Nice
Mediterranean Sea
Spain
(b)
Montpellier
Marseilles
Nice
Shoreline
generalized
(c)
Montpellier
Marseilles
Nice
(d)
Nice
Marseilles
Montpellier

Principle #11
(Influence of neighbors)

- ***"In geographic repositories, do not forget that objects at the vicinity (outside the jurisdiction) can have an influence".***

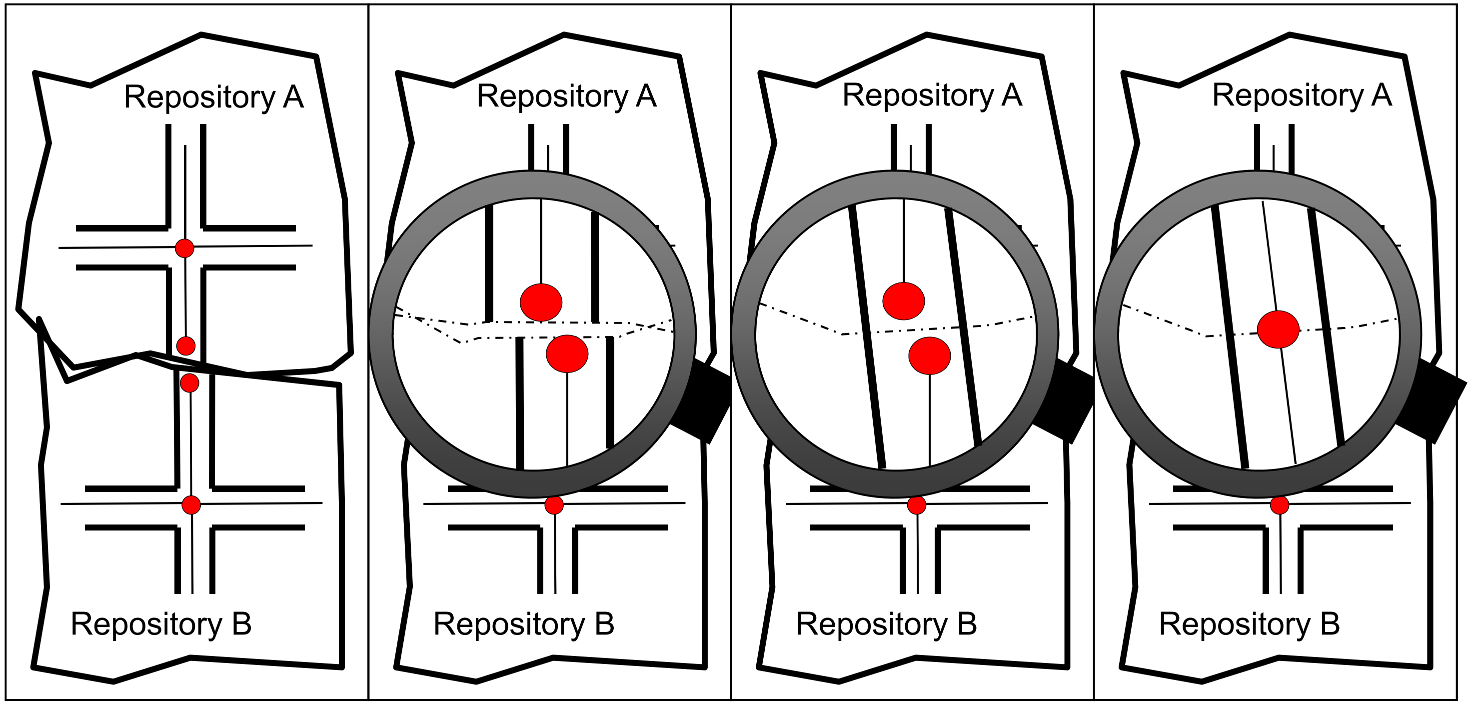
- Ex. Geneva and French Region Rhône-Alpes

Principle #12
(Cross-boundary interoperability)

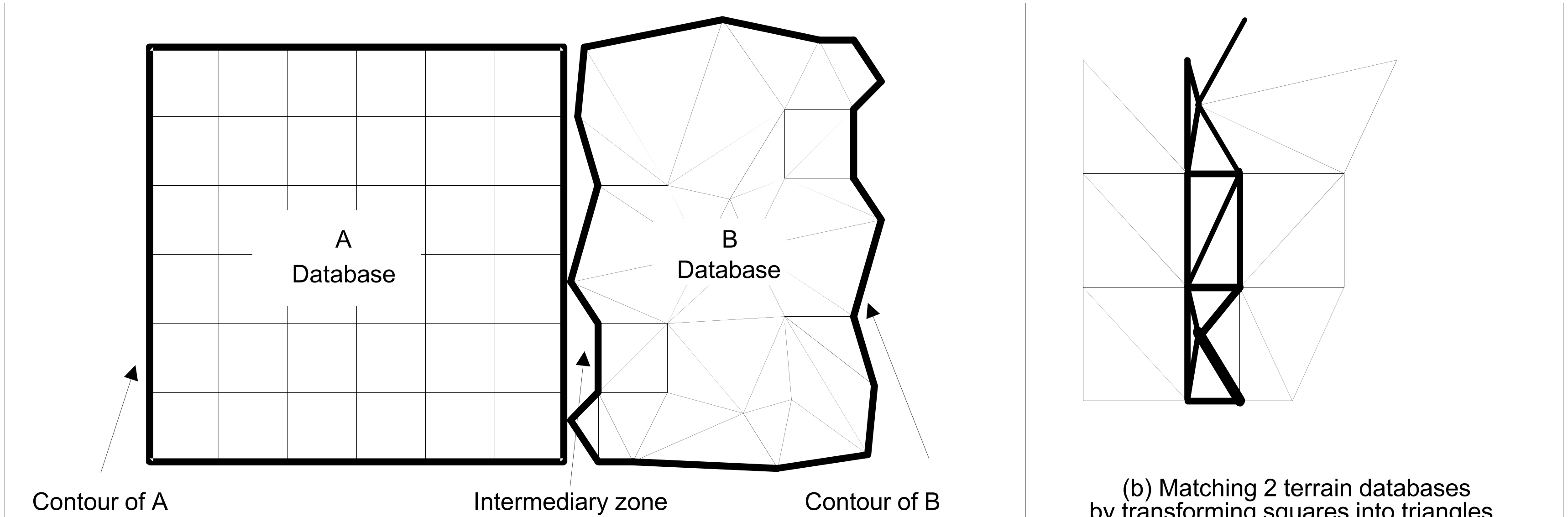
- ***"Any geographic repository must provide key-information to ensure cross-boundary interoperability".***

- Two cases:
  - Network continuity
  - Terrain continuity

Road continuity



Terrain continuity

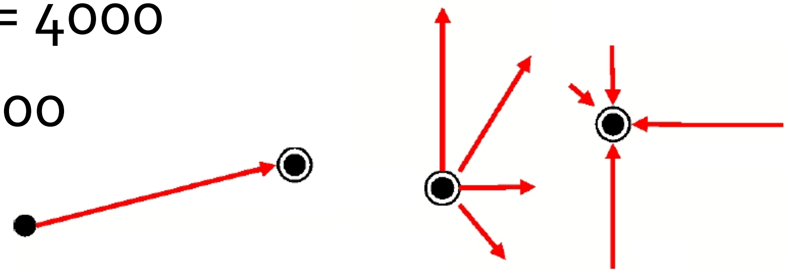


(a) Two adjacent terrain databases

(b) Matching 2 terrain databases by transforming squares into triangles and adding some intermediary triangles

Elementary knowledge (1/2)

- Facts
  - Spain.population= 30 000 000
  - Spain.geometry = { ******}
- Flow
  - Bi-directional flow
    - Flow (Barcelona, Madrid) = 4000
    - Flow (Sevilla, Valencia) = 300
  - Converging flows
  - Diverging flows



Elementary knowledge (2/2)

- Clusters
  - UK= Union (England, Scotland, Wales, NorthenIreland, etc)
- Geographic relations
  - Location rules
  - Topological relations
  - The more of this/the more of that
  - Co-location (CityHall, Church)

Where to find GK?

- Discussions with experts
- Spatial data mining
- Analyzing web documents
  - Gazetteers
  - Ontologies

5 – Conclusion (1/2)

- Importance of geographic knowledge
- Several layers
  - Generic layers
  - Specific layers
  - Application layers
- First steps to geographic reasoning

Conclusions (2/2)

- Other minor contributions
  - Ribbon
  - Ribbon topology
  - Homology relations
  - Generalization of topological relations
  - Visual knowledge representation

Main recent references

- LAURINI R. (2014) "*A Conceptual Framework for Geographic Knowledge Engineering*", Journal of Visual Languages and Computing, Volume 25, pp.2-19.

- LEJDEL B., KAZAR O., LAURINI R. (2015). "*Mathematical framework for topological relationships between ribbons and regions*", Journal of Visual Languages & Computing. Volume 26, pp. 66-81.

- LAURINI R. (2015) "*Geographic Ontologies, Gazetteers and Multilingualism*" Journal Future Internet, January 2015.

- LAURINI R. (2015) "*Fundamentals of Geographic Knowledge Engineering for Territorial Intelligence*" in the book "Knowledge Engineering: Principles, Methods and Applications" To be published by for NovaPublishers.

To download this talk:

http://liris.insa-lyon.fr/robert.laurini/ftp/GKS.zip

For any contact:

Robert.Laurini@insa-lyon.fr

Thanks for your attention!