MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Spatial Database Systems

Karine Reis Ferreira
karine@dpi.inpe.br

Aula da disciplina Computação Aplicada II (CAP 236) – 2015

Geographic Information System (GIS)

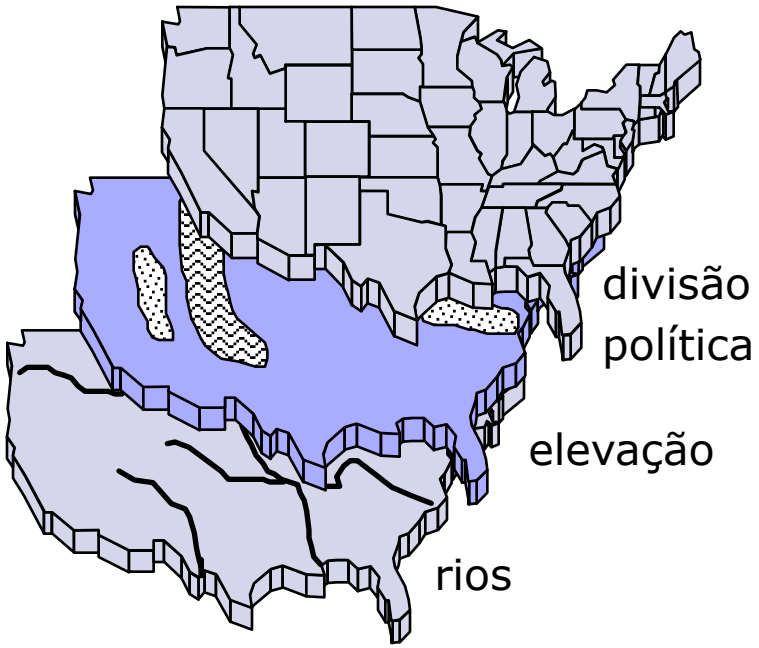
- *Sistemas de Informações Geográficas* (*SIG*)

- Computational system able to store, process and handle geographical data (Worboys and Duckham, 2004)

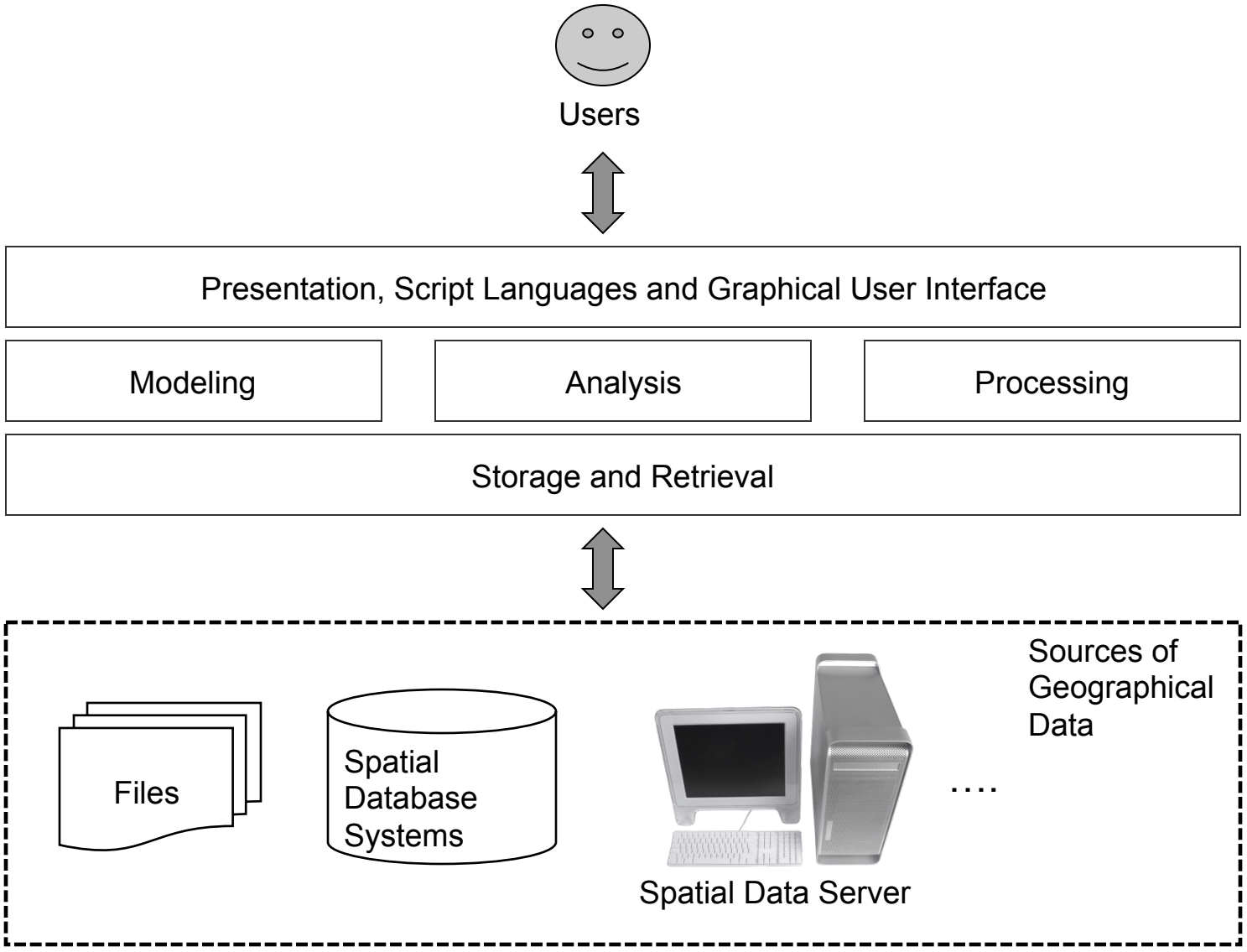
- GIS functions (Rigaux et al, 2002):
  - Data input and verification;
  - Data storage and management;
  - Data output and presentation;
  - Data transformation;
  - Interaction with end users;

GIS: Data Sets are organized in Layers

- Layer (nível, camada, plano de informação):
  - Contains information associated to a specific theme.
  - Spatial extent
  - Examples:
    - Divisão política
    - Elevação
    - Rios

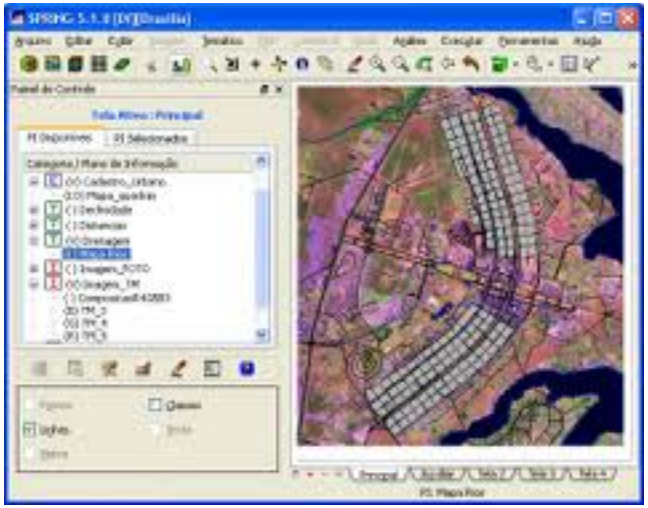


Geographic Information System (GIS)

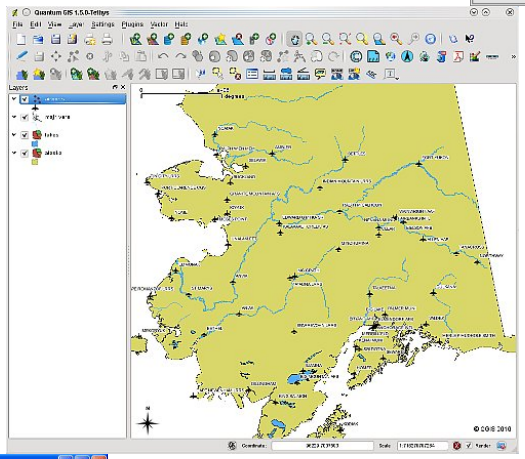


Examples of GISs

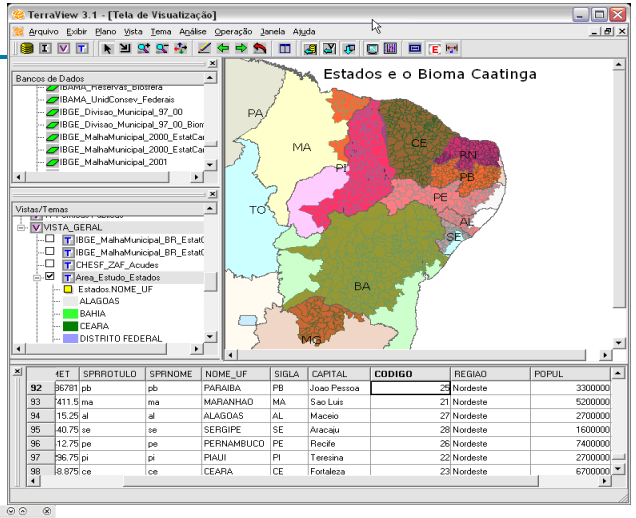
SPRING



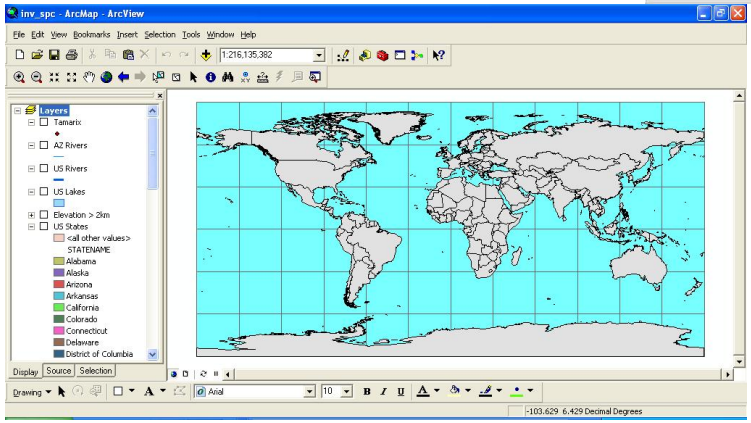
Quantum GIS



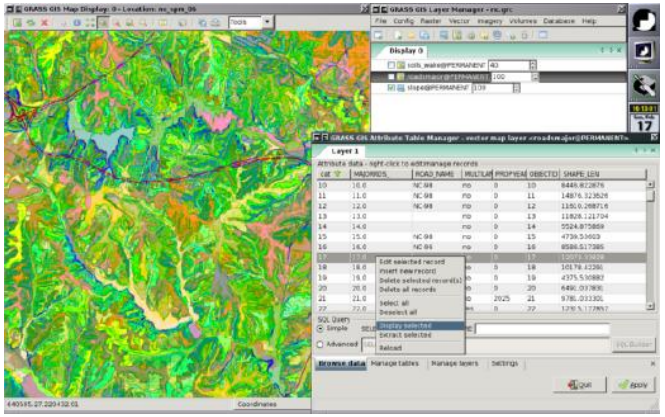
TerraView



ArcGIS

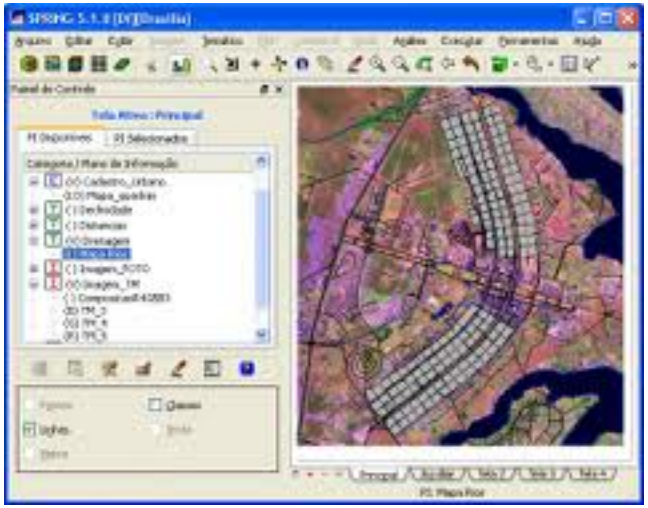


GRASS

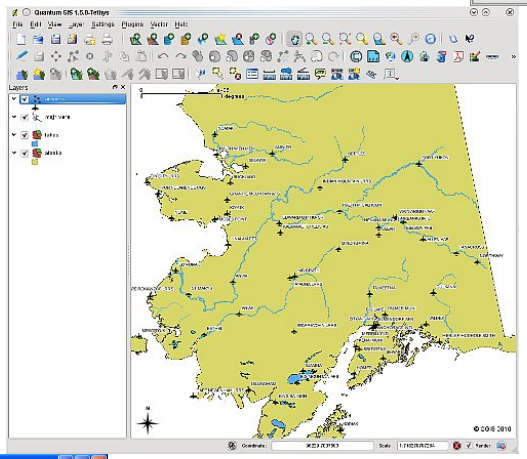


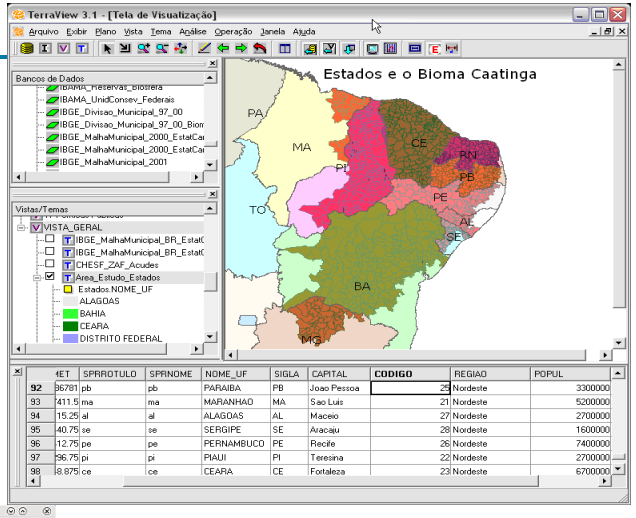
Examples of GISs

SPRING → Free



Quantum GIS



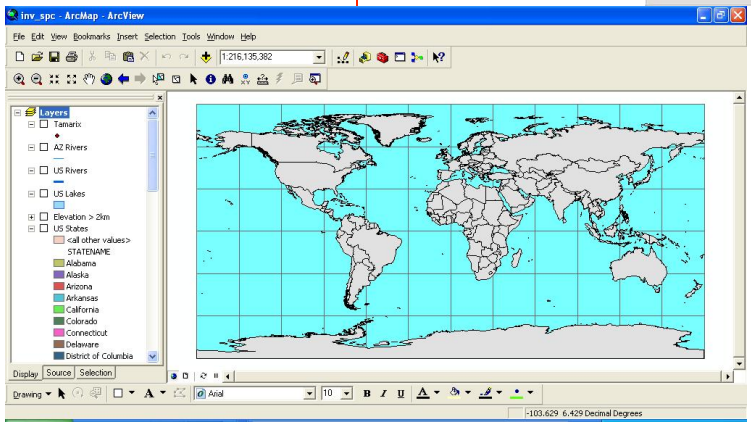


TerraView

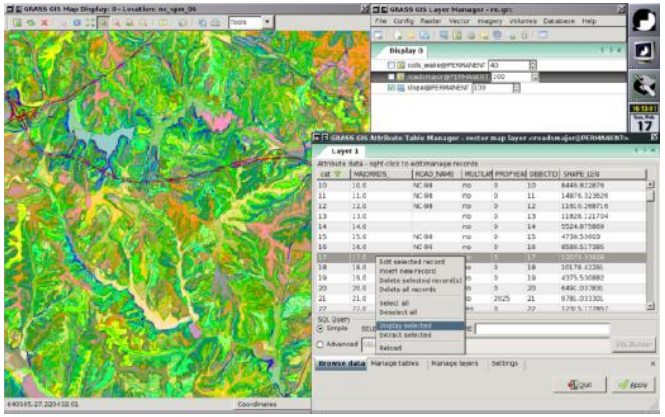
Free and Open Source

Commercial

ArcGIS



GRASS



Geographic Information System (GIS)

- *Sistemas de Informações Geográficas* (*SIG*)

- Computational system able to store, process and handle geographical data (Worboys and Duckham, 2004)

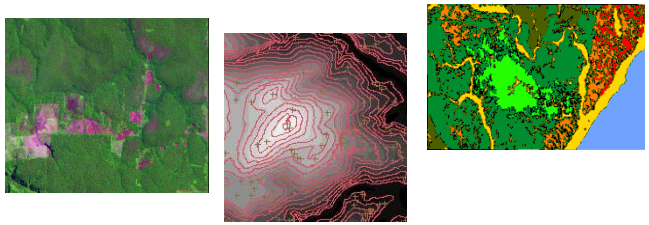
- GIS functions (Rigaux et al, 2002):
  - Data input and verification;
  - Data storage and management;
  - Data output and presentation;
  - Data transformation;
  - Interaction with end users;

Representation of Geographical Data

Conceptual Models

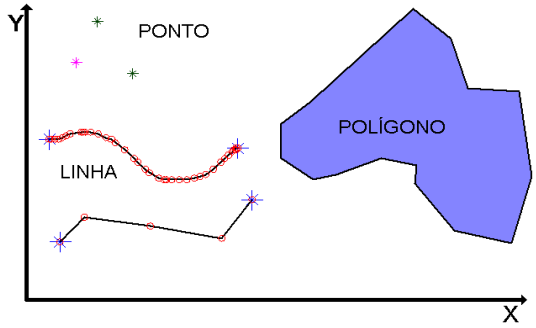
(1) Object-based models
(entity- or feature-based models)

(2) Field-based models
(space- or coverage-based models)

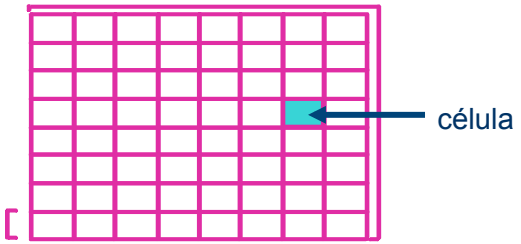


Computational Representation

Vector Model

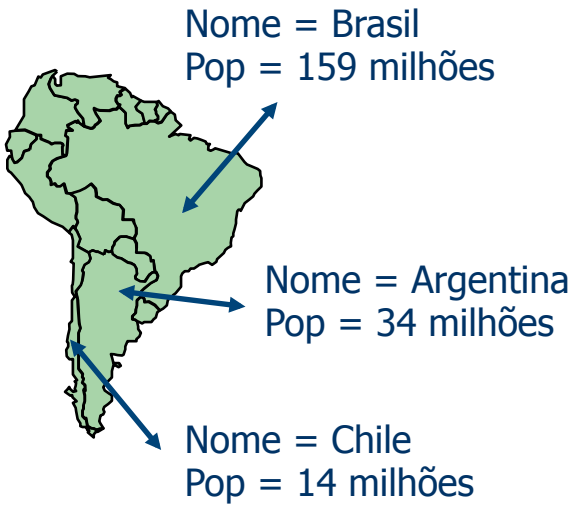


Tessellation Model



Geographical Objects (Entities, features, …)

- An object has two components: (a) a description and (2) a spatial component (spatial extent)
- Well-defined boundary
- Identity
- Internal homogeneity



Fields (Coverages, surfaces, …)

- Each point in space is associated one or several attribute values, defined as continuous functions in x and y.
- Continuous phenomena collected as attribute values varying with the location in the plane. Example: precipitation, temperature, and pollution.

IDH

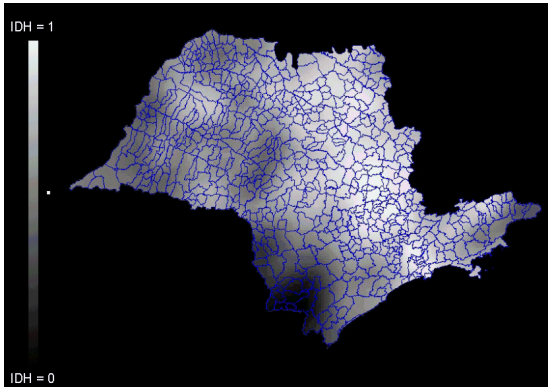
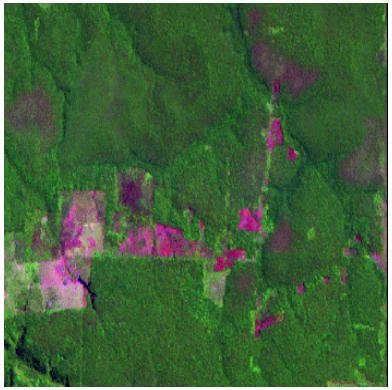
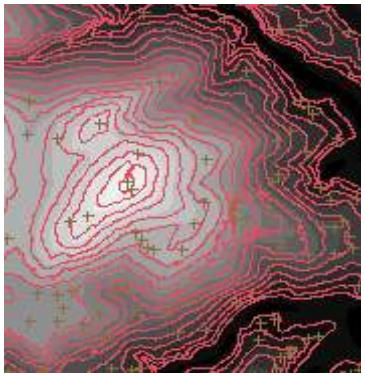


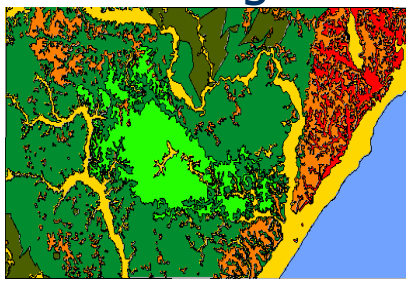
Imagem satélite



Altimetria



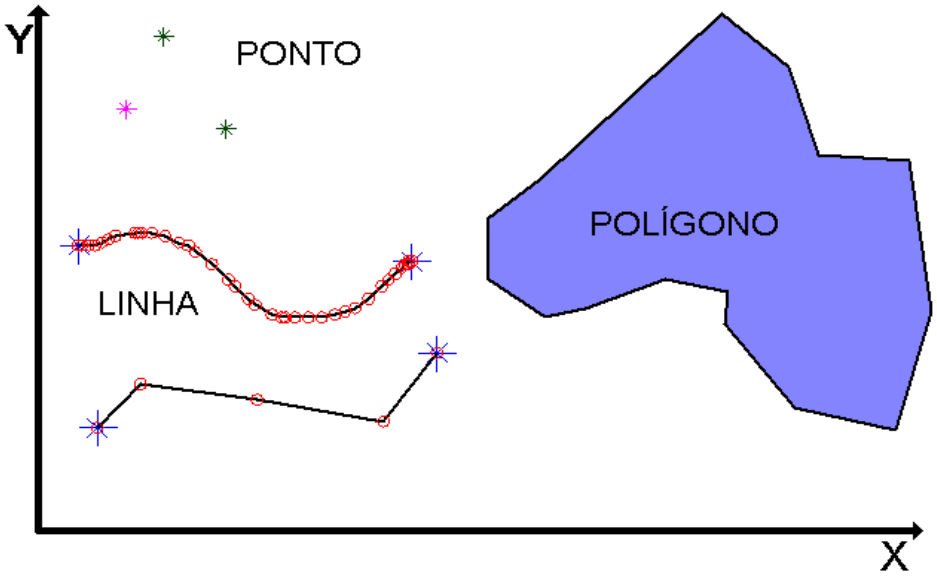
Geologia



Vector Model

Objects are constructed from points and edges as primitives.

- *Zero-dimensional* objects or *points*
- *One-dimensional* objects or *linear objects (polyline)*
- *Two-dimensional* objects or *surface objects (polygon)*



Vector Data Types

Source: Rigaux et al (2002)

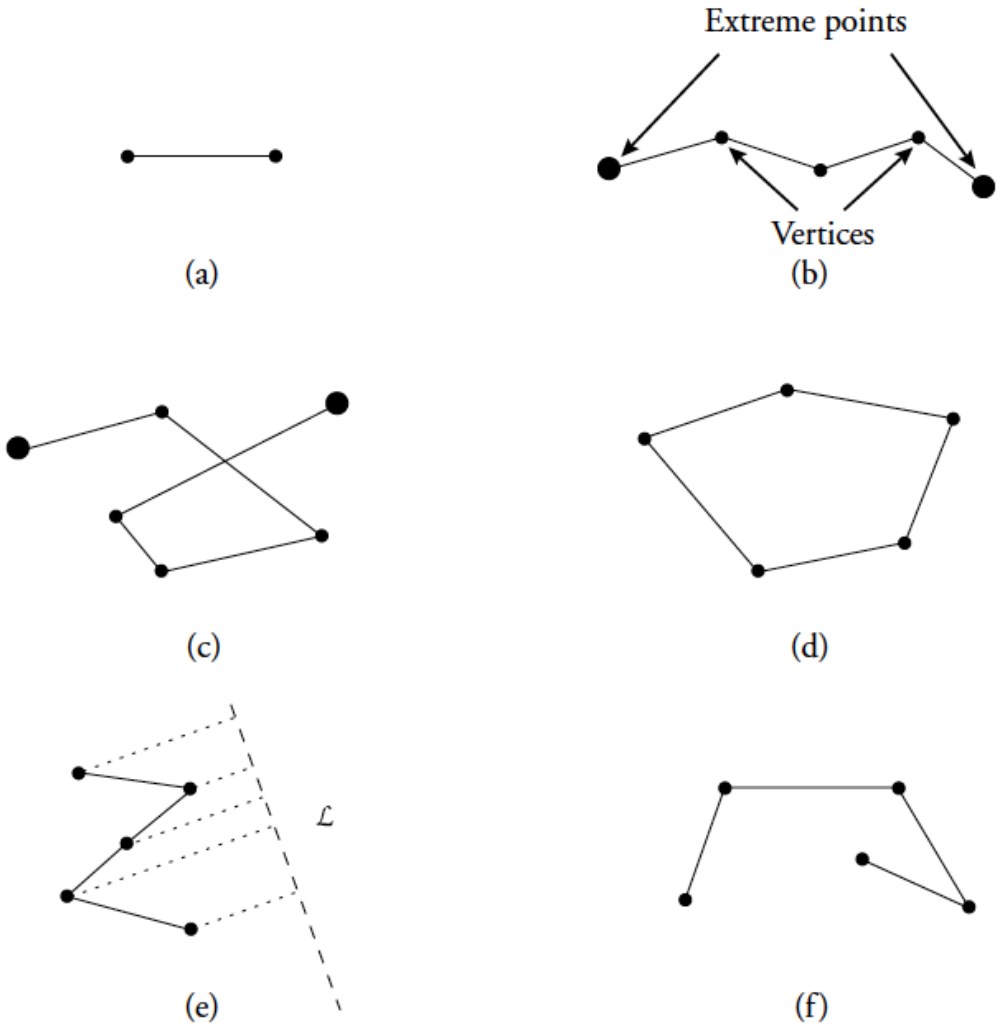


Figure 2.1 Examples of one-dimensional objects: line segment (edge) (a), polyline (b), non-simple polyline (c), simple closed polyline (d), monotone polyline (e), and non-monotone polyline (f).

Vector Data Types

Source: Rigaux et al (2002)

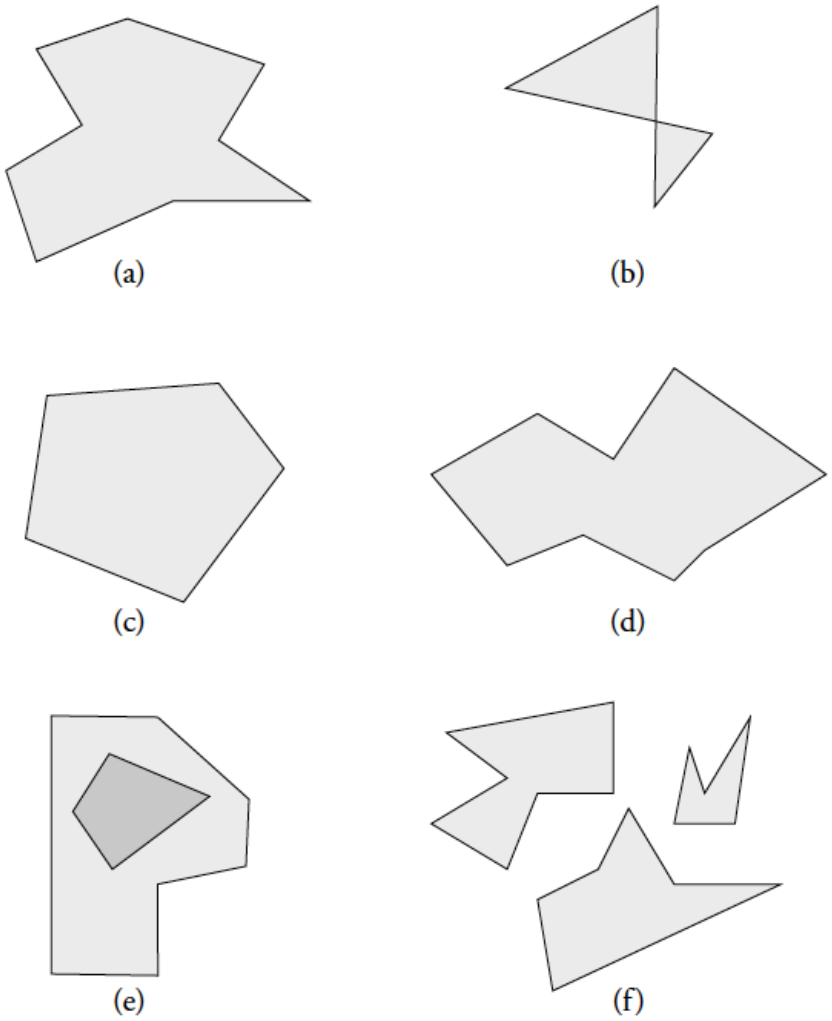


Figure 2.2 Examples of 2D objects: simple polygon (a), non-simple polygon (b), convex polygon (c), monotone polygon (d), polygon with hole (e), and region (f).

Vector Data Types

Source: Rigaux et al (2002)

- *point : [x: real, y: real]*
- *polyline : < point >*
- *polygon : < point >*
- *region : { polygon }*

< [4, 4], [6, 1], [3, 0], [0, 2], [2, 2] >

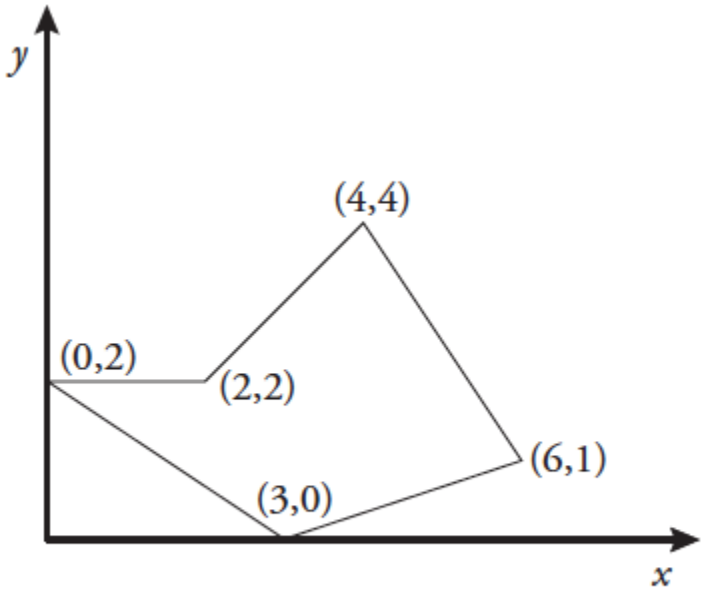


Figure 2.6 Vector representation of polygon *P*.

Tessellation Models

Can be divided into *fixed* (or *regular*) and *variable* (or *irregular*) tessellation modes.

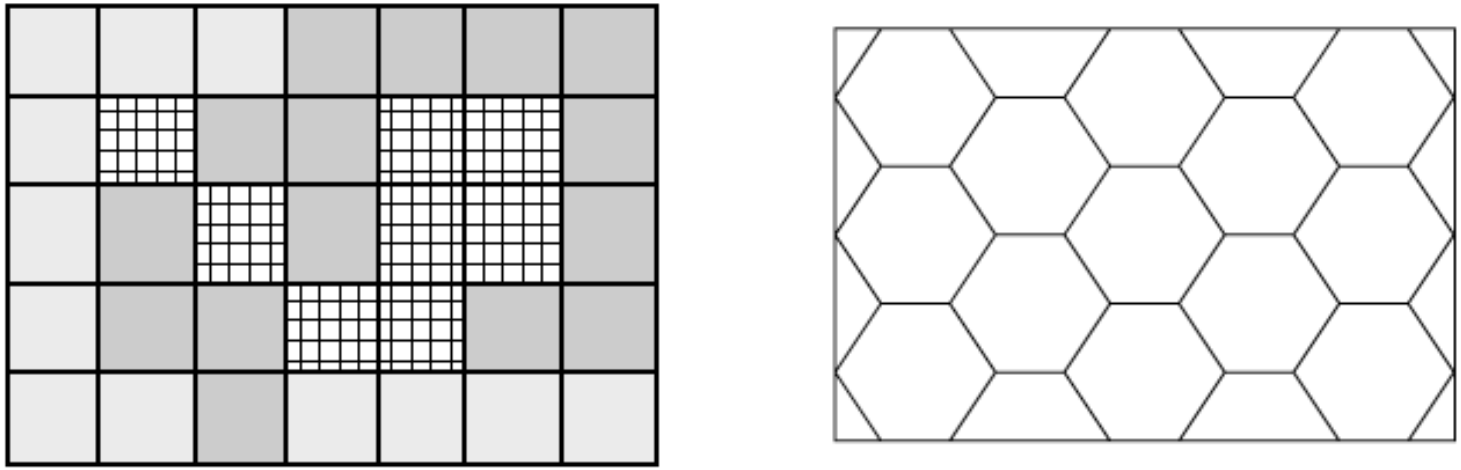


Figure 2.3 Regular tessellations: grid squares (a) and hexagonal cells (b).

Source: Rigaux et al (2002)

Tessellation Models

Can be divided into *fixed* (or *regular*) and *variable* (or *irregular*) tessellation modes.

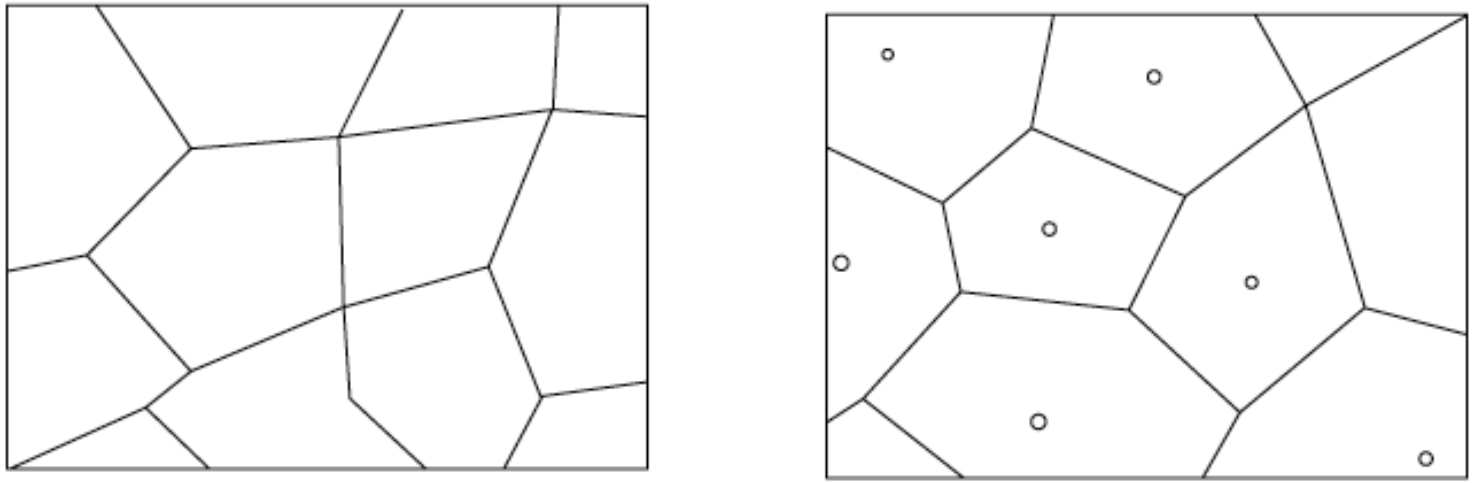


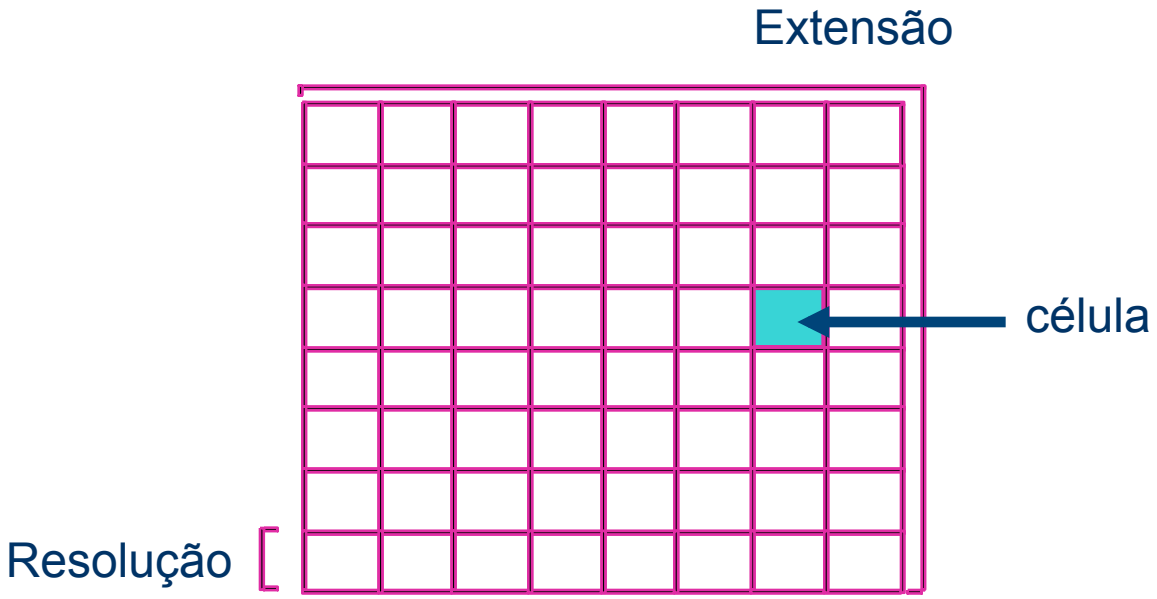
Figure 2.4 Irregular tessellations: cadastral zones (a) and Thiessen polygons (b).

Source: Rigaux et al (2002)

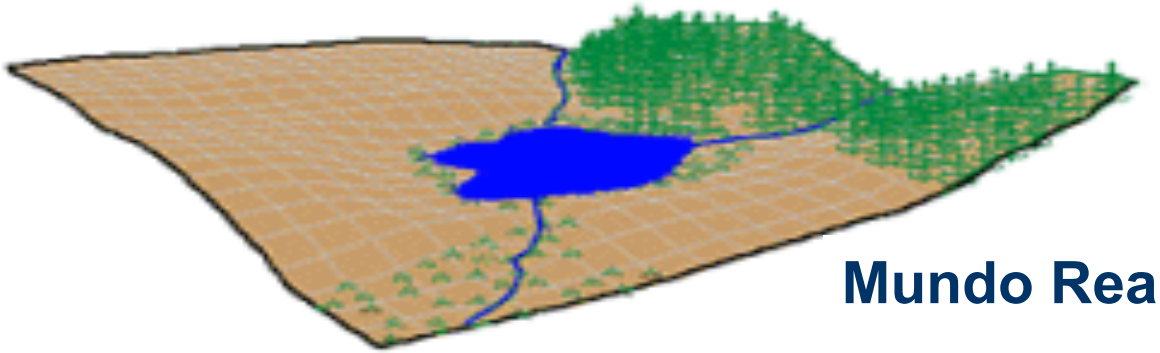
Raster data type

A fixed representation uses a regular grid or *raster*, which is a collection of polygonal units of equal size:

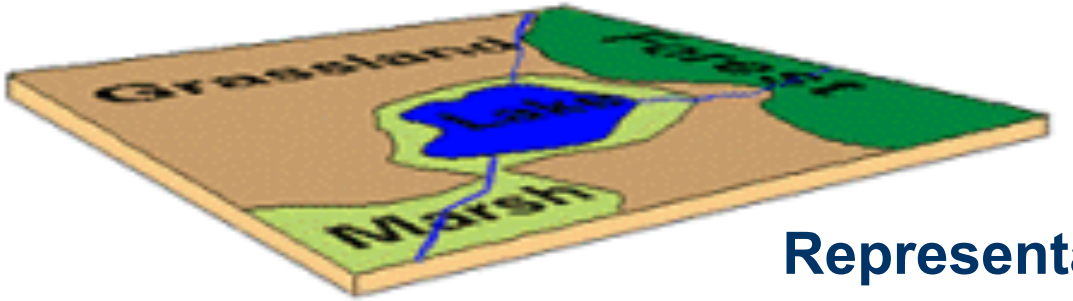
- ✓ finite number of elementary cells;
- ✓ regular 2D grid of N × M rectangular cells;
- ✓ the cells are called *pixels*;
- ✓ a pixel has an address in the plane: (i, j)



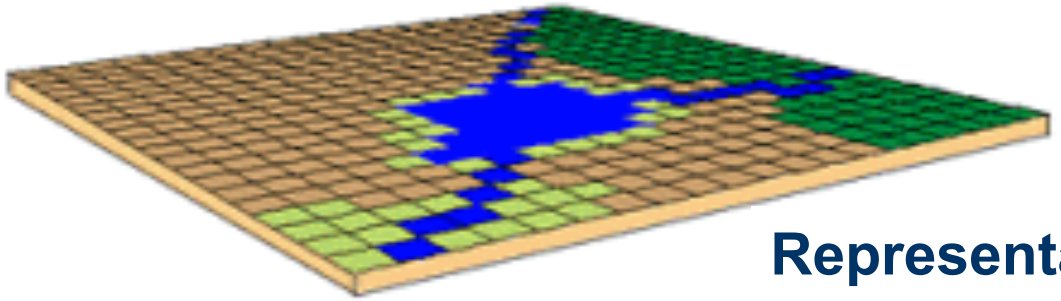
Vector X Raster Representation



Mundo Real



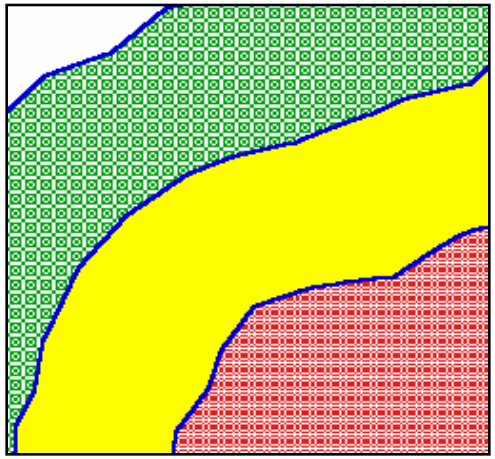
Representação Vetorial



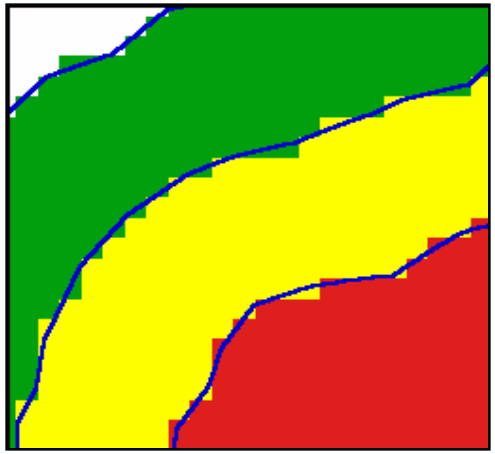
Representação Matricial

Vector X Raster Representation

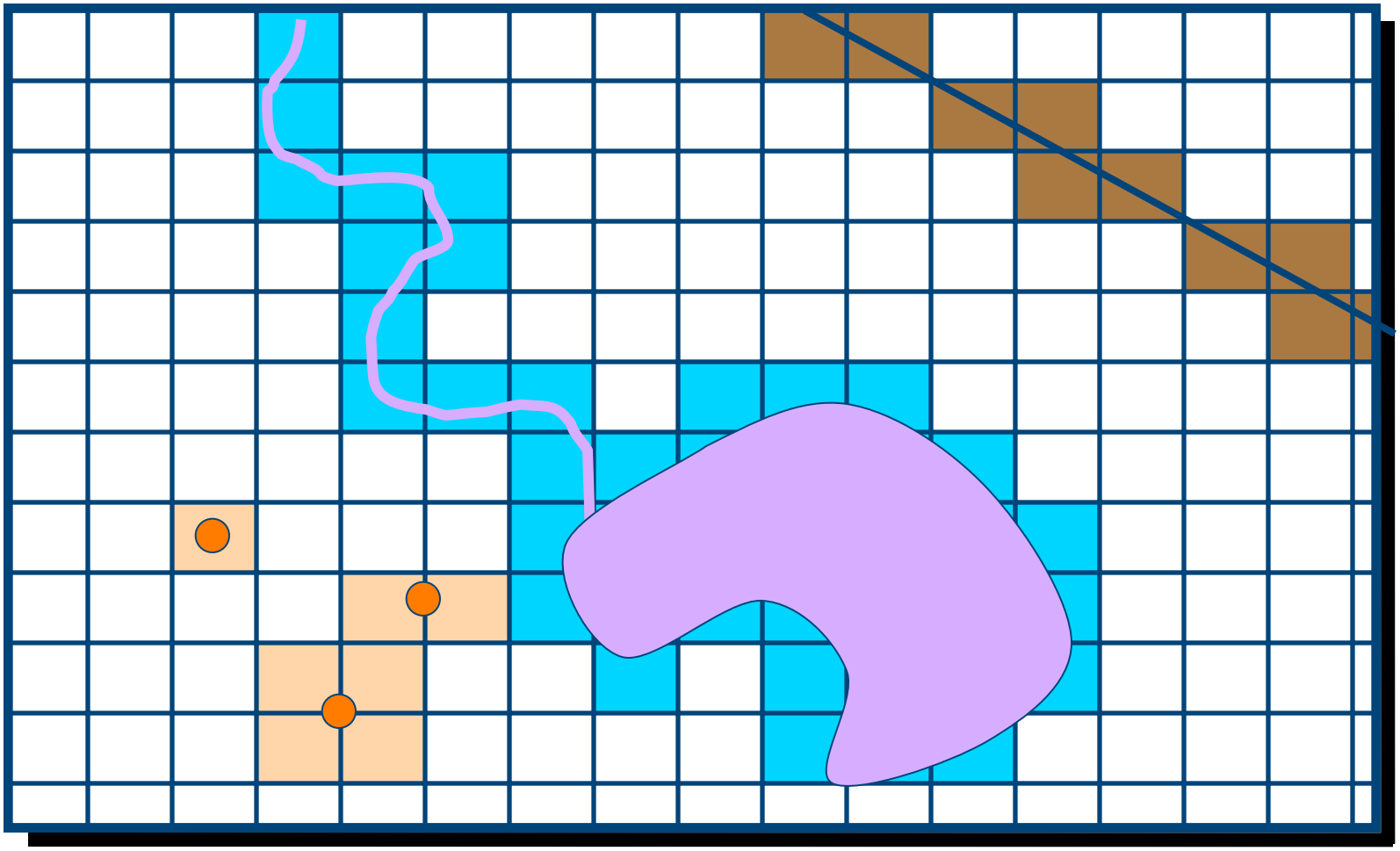
- Vector
  - Objects with well-defined boundaries
  - Topological relationships
  - Attributes associated to spatial elements
  - Efficient storage



- Raster
  - Continuous phenomena
  - Attribute values varying within a spatial extent
  - Each point has a attribute value associated

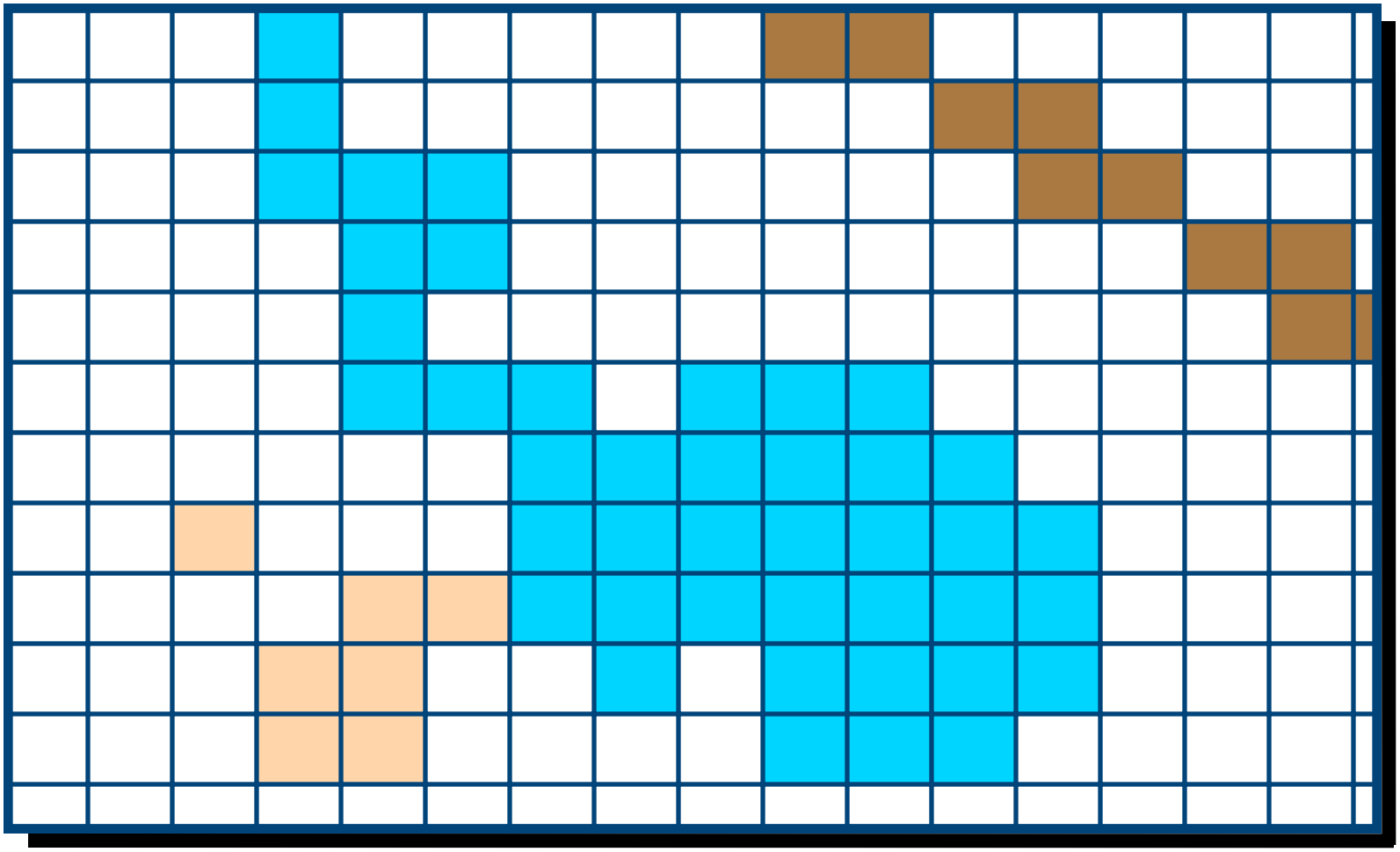


Vector => Raster



fonte: Mohamed Yagoub

Vector => Raster



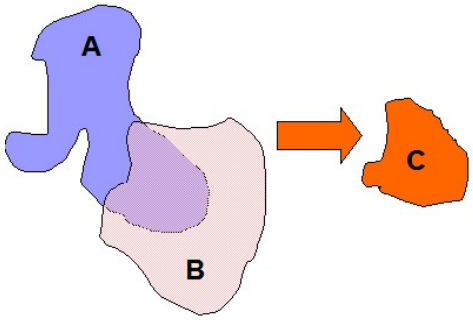
fonte: Mohamed Yagoub

GIS: Typical Operations

- Theme selection:
  - Example: *Name and population of countries of 50 million inhabitants or more.*
- Metric operations: area, perimeter, length, distance, …
  - Example: *What is the distance between Paris and Berlin?*
- Topological operations: related to the (topological) relationships existing among data.
  - Example: *What are the countries adjacent to Belgium? What cities can I reach by train from Berlin without any stop?*

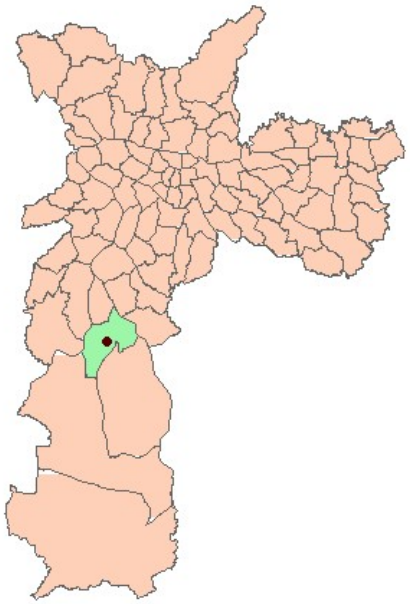
GIS: Typical Operations

- Operation that creates new geometries:
  - Buffer, centroid, convex hull, etc.
- Set operations:
  - Intersection, union, difference, etc.

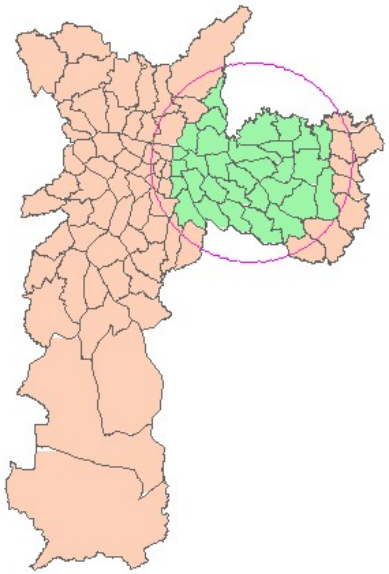


Geometric selection

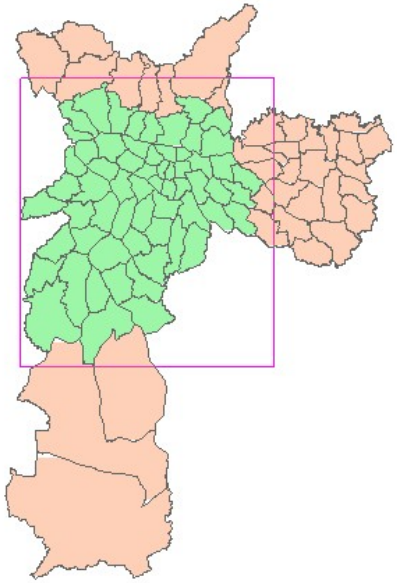
- *Window query*: returns objects that overlap a given area or *window*, which is usually rectangular.
- *Point query*: retrieves all objects whose geometry contains a given point.



Point query



Window query



Window query

Fonte: Karine Ferreira (2006)

Geometric selection

Clipping: extracts the portion of a theme located within a given area. The result corresponds exactly to the intersection of the geometry of the geographic objects and the geometry of the area.

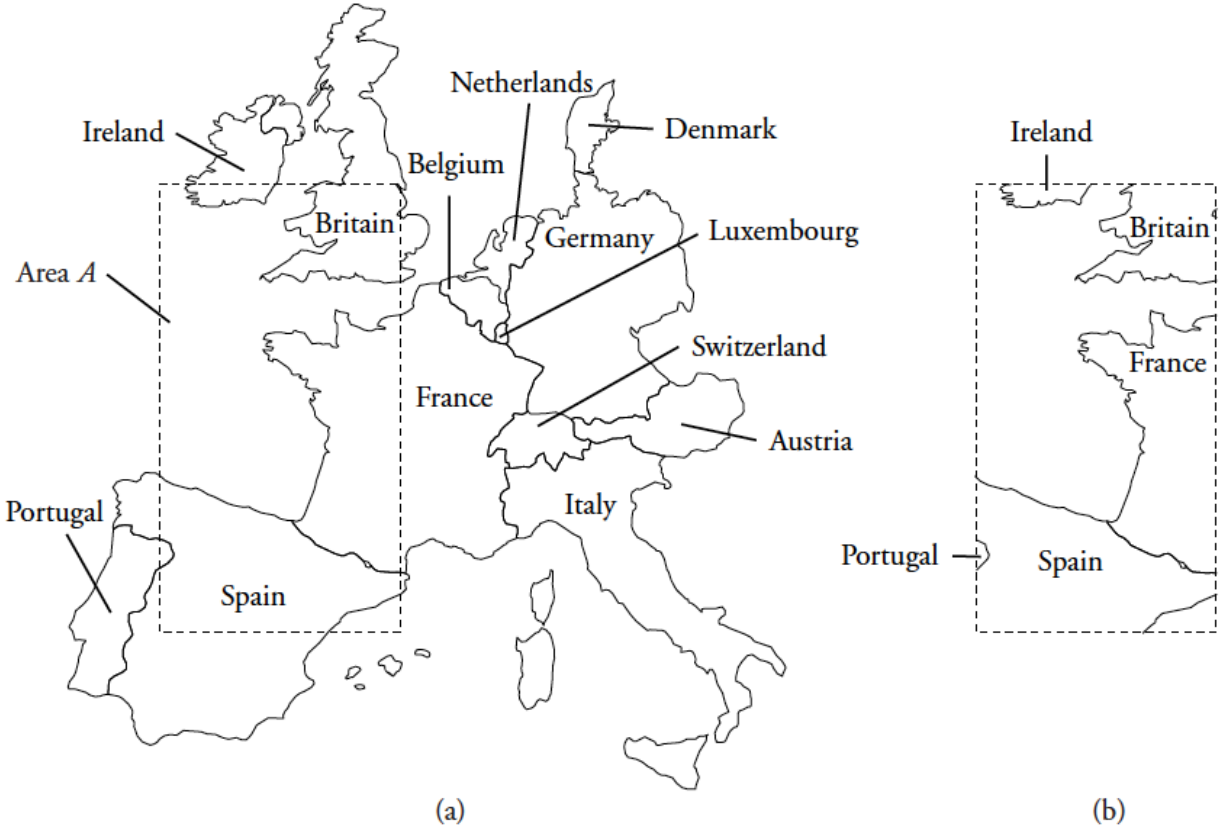


Figure 1.8 Clipping: theme T (a) and clipping of T with area a (b).

Source: Rigaux et al (2002)

Merger or Aggregation Operation

Performs the geometric union of the spatial part of *n* geographic objects that belong to the same theme, under a condition supplied by the end user.

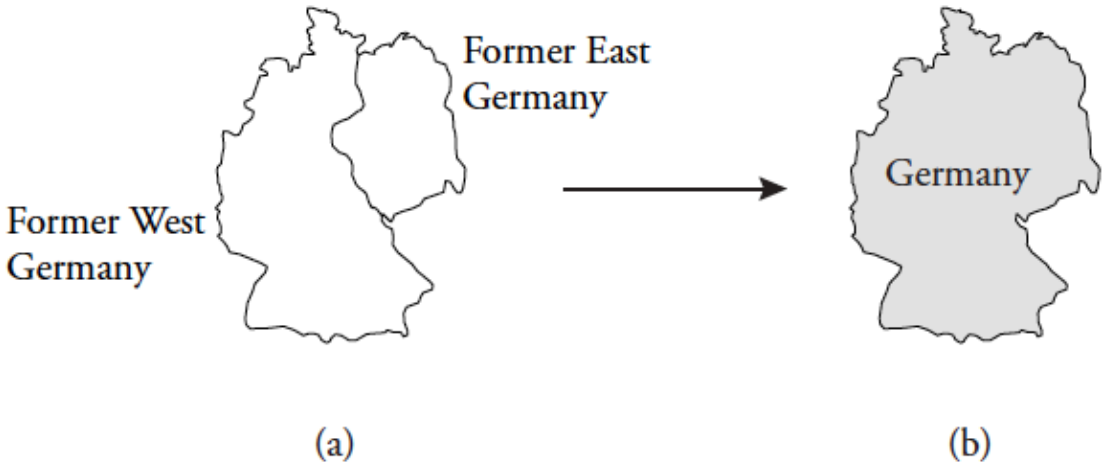


Figure 1.9 Merging two geographic objects in a theme: theme with two geographic objects (a) and theme with one geographic object (b).

Source: Rigaux et al (2002)

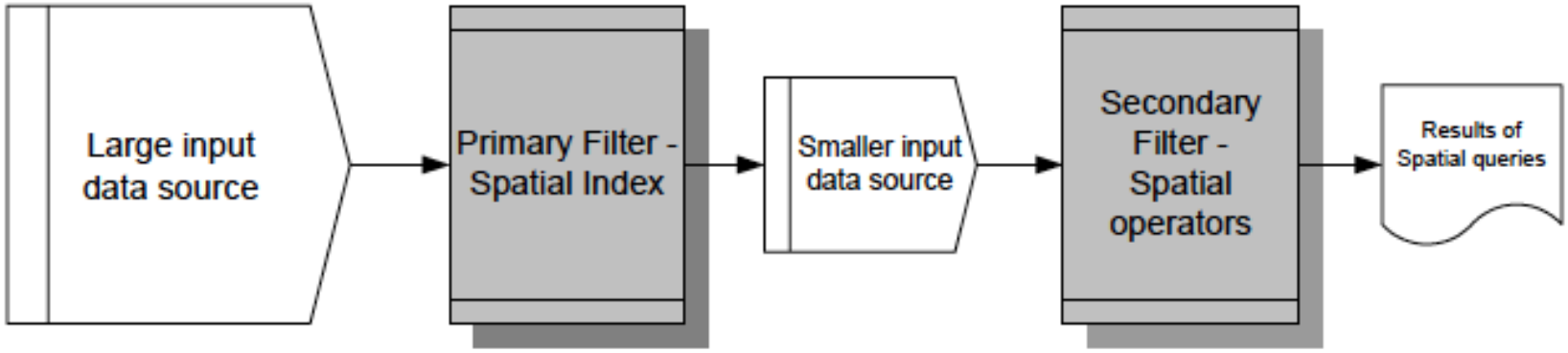
Spatial Join

- Given two sets of geographical objects D_1 and D_2 and a spatial predicate θ, determine all pairs $(d_1, d_2) \in D_1 \times D_2$ whose geometries satisfy θ.

Example: For each street, returns the schools that are less than 1000 meters from it.

Fonte: Karine Ferreira (2006)

Spatial Operations – Two Tiers



Spatial Index

- Spatial Access Methods (SAMs)
- Most of them are based on a simple approximation of the geometry (minimal bounding box - mbb)

Aproximação pelo ***Retângulo Envolvente Mínimo***
(REM ou Bounding Box ou MBR)

Fixed Grid

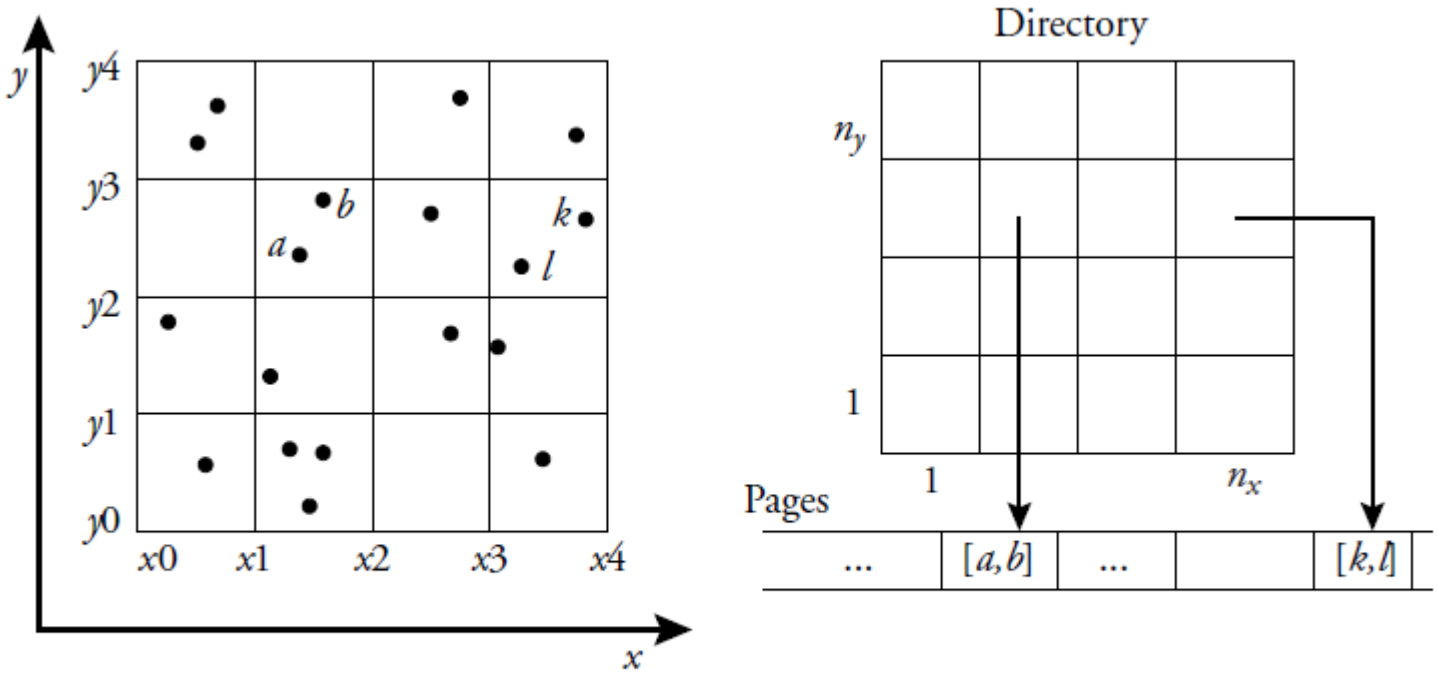


Figure 6.2 Fixed grid.

Fonte: Rigaux et al (2002)

Fixed Grid

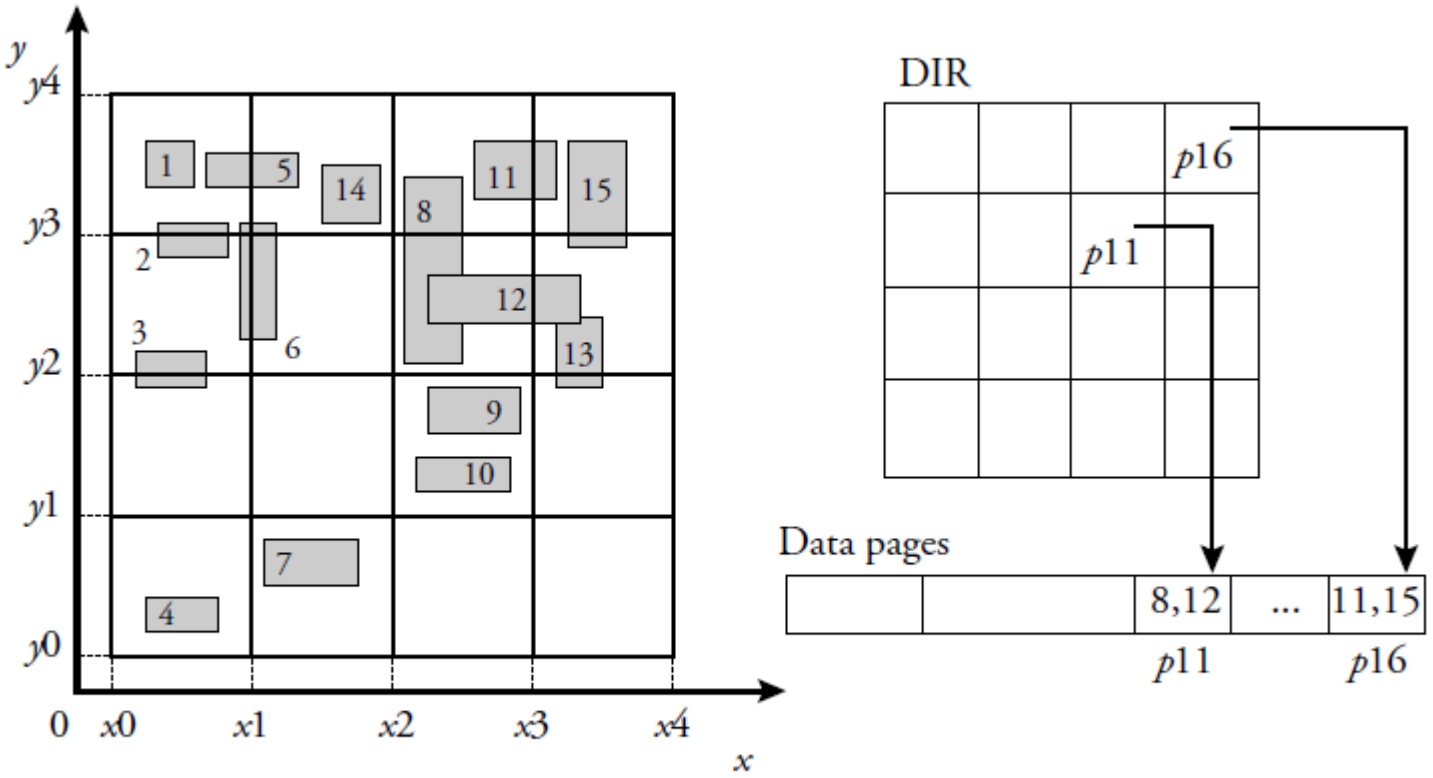


Figure 6.5 A fixed grid for rectangle indexing.

Fonte: Rigaux et al (2002)

QuadTree

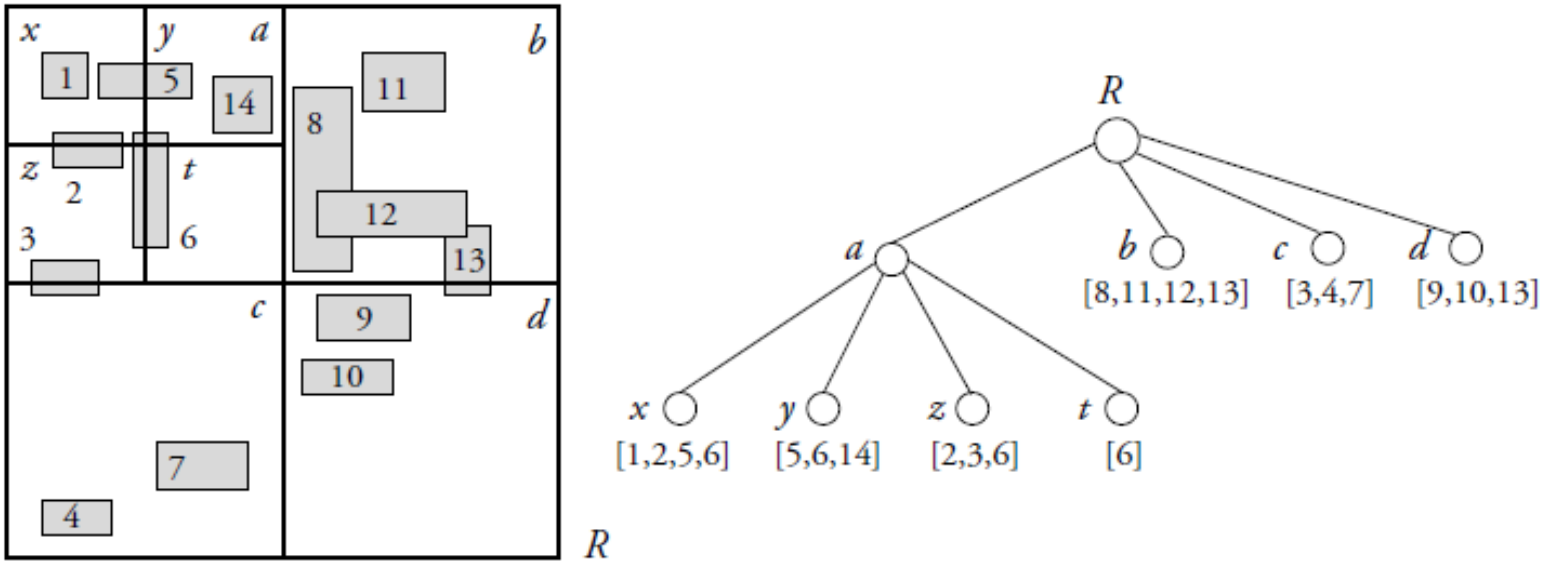


Figure 6.10 A quadtree.

- Árvore: cada “nó” ou “tronco” gera quatro “folhas”
- Cada nó corresponde a uma região quadrada do espaço
- Cada região é subdividida em quatro partes iguais sucessivamente até ter um ou nenhum objeto geográfico dentro de cada quadrante.

Fonte: Rigaux et al (2002)

QuadTree

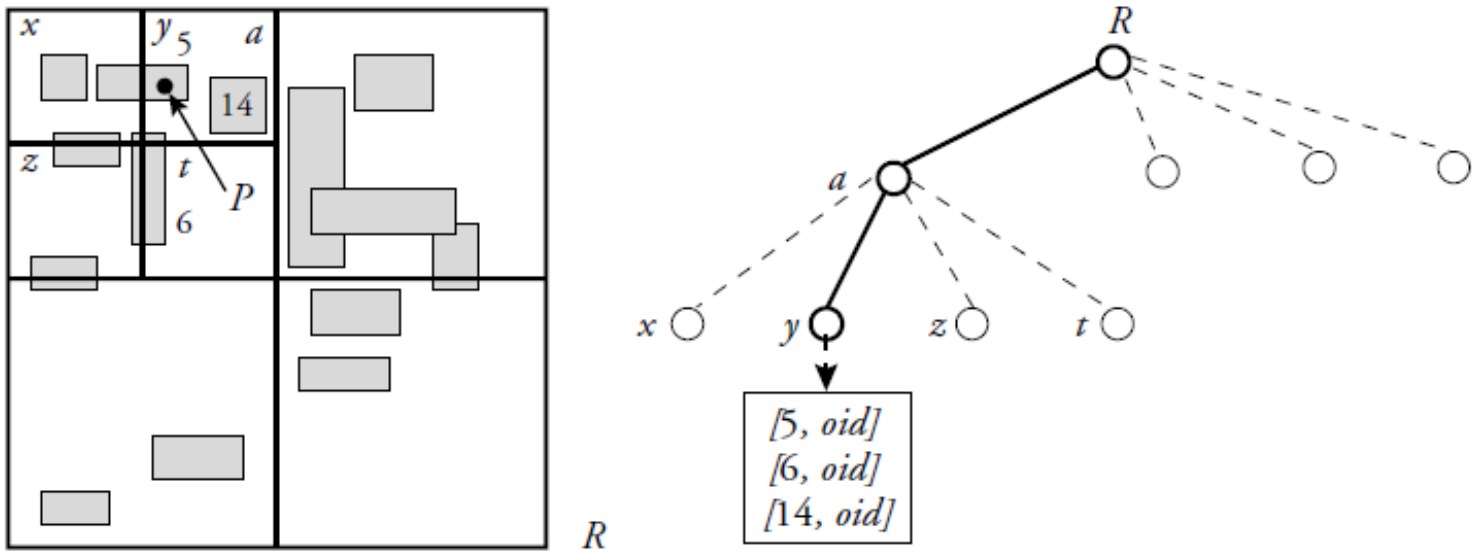


Figure 6.11 A point query on a quadtree.

Fonte: Rigaux et al (2002)

RTree

- Árvore-R

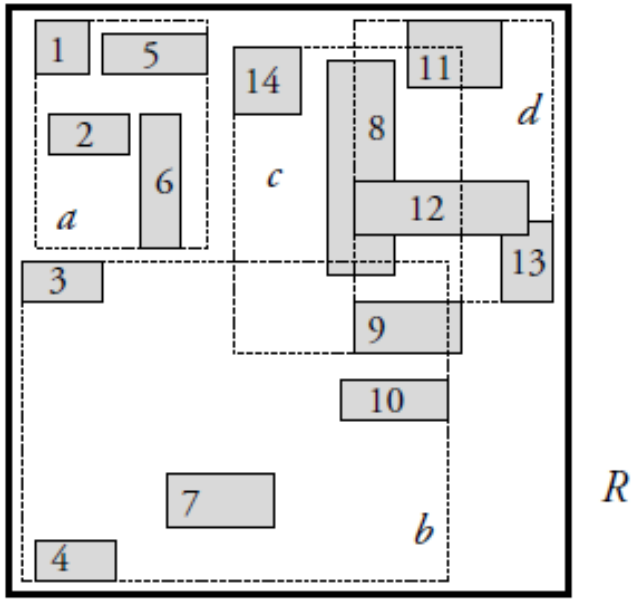
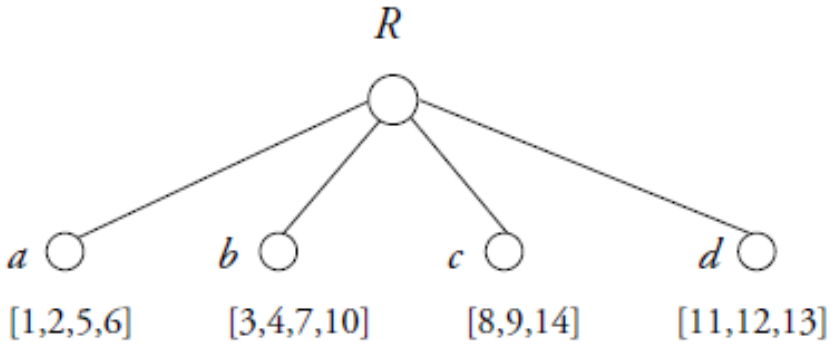
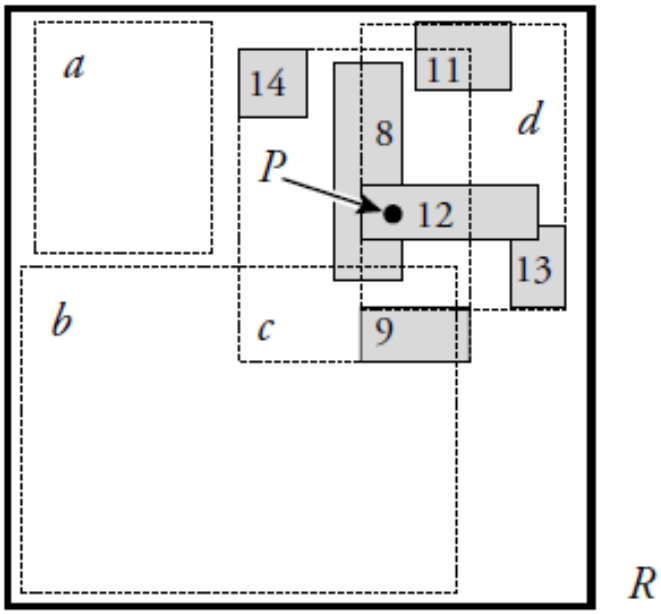


Figure 6.22 An R-tree.



Fonte: Rigaux et al (2002)

RTree



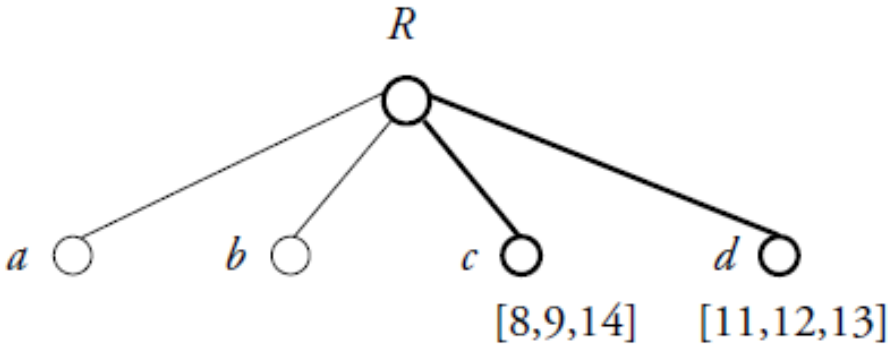


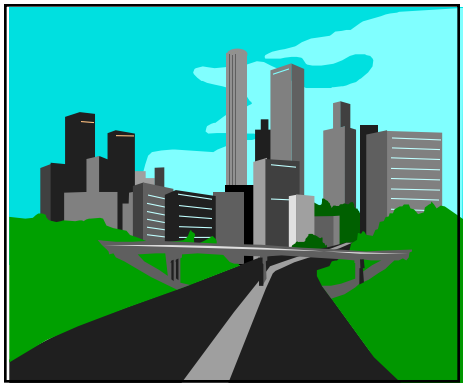
Figure 6.24 Point queries with R-trees.

Fonte: Rigaux et al (2002)

Evolução dos SIGs

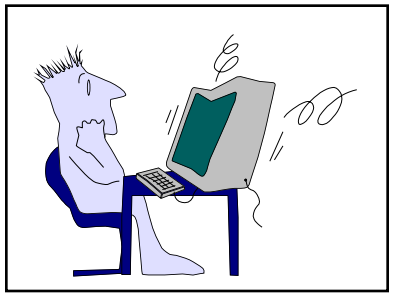
(1) Gerenciamento e utilização dos dados geográficos

Servidores WEB



SIG Distribuído
(multiusuários)

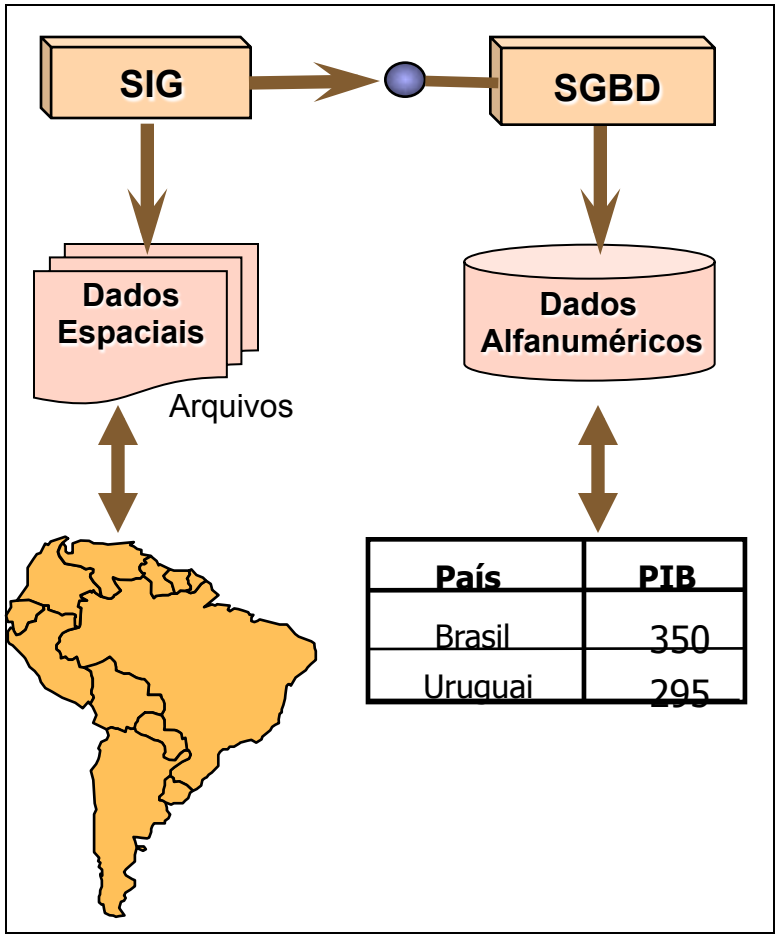
SIG "DeskTop"



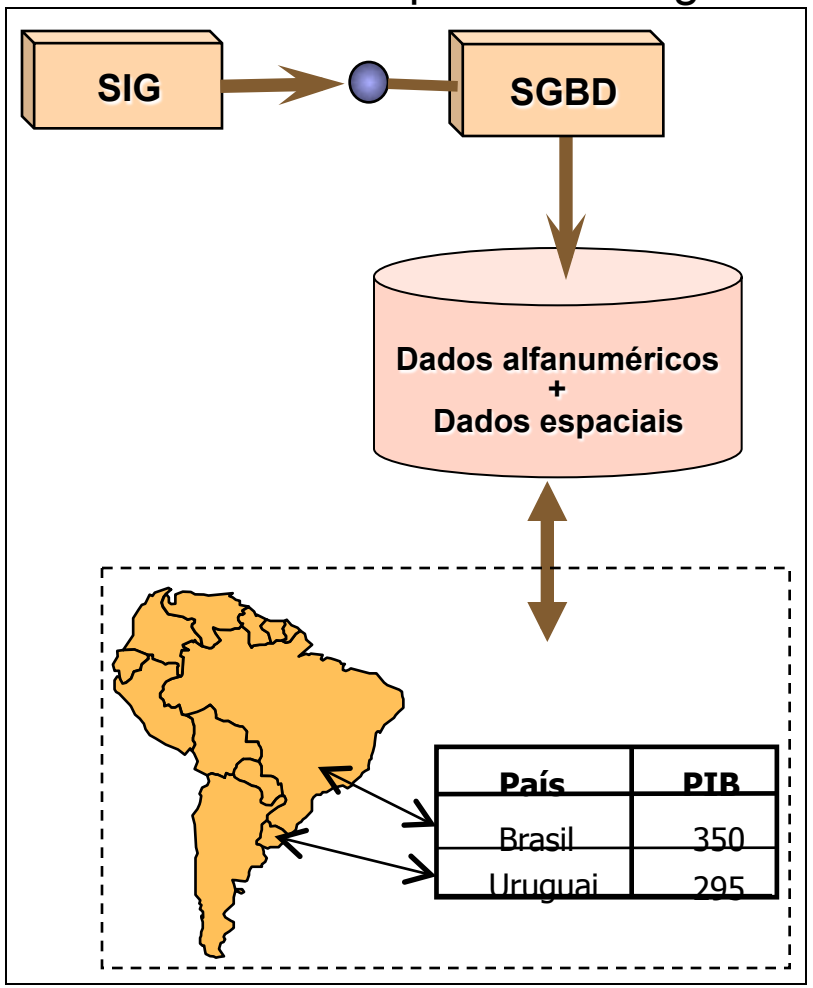
Evolução dos SIGs

(2) Armazenamento de dados geográficos

Arquitetura Dual



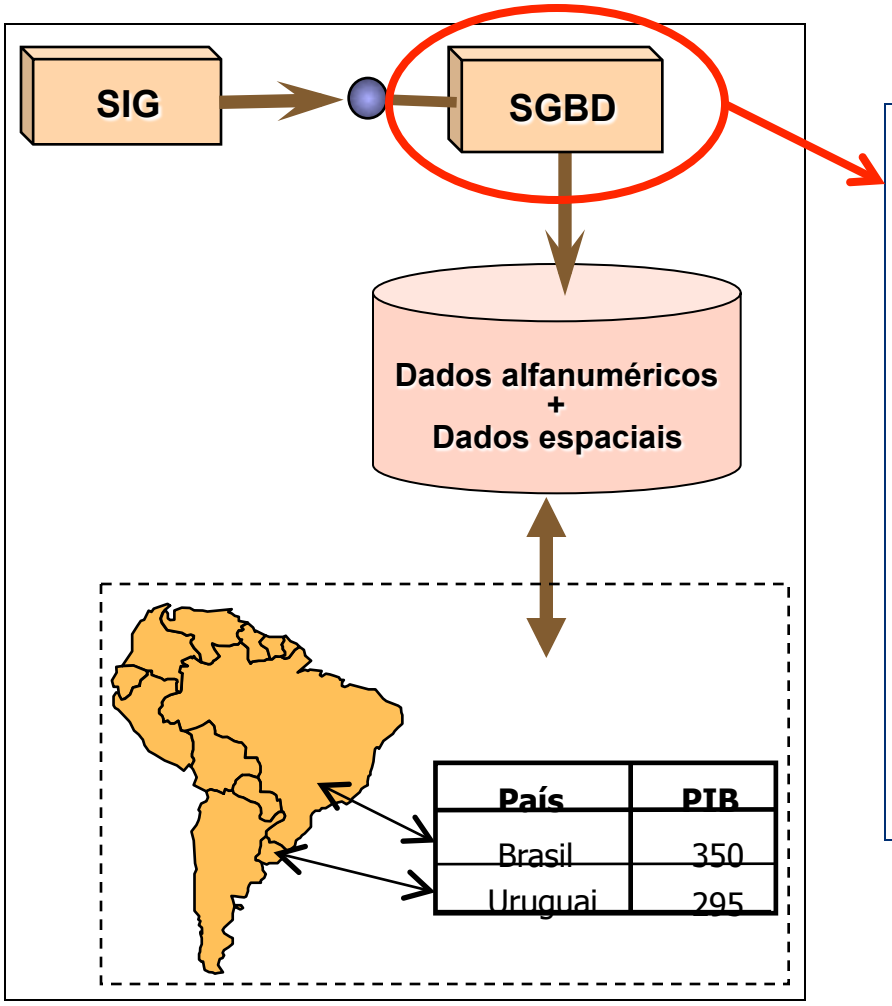
Arquitetura Integrada



SIGs e SGBDs

- Os Sistemas de Informações Geográficas (SIG) estão evoluindo para utilizar Sistemas Gerenciadores de Bancos de Dados (SGBD) para armazenar e gerenciar dados geográficos:
  - A responsabilidade de gerenciamento dos dados geográficos passa a ser dos SGBDs e não dos SIGs.
  - SGBDs tradicionais são estendidos para suportar dados espaciais.

SIGs e SGBDs



Os SGBDs objeto-relacionais são estendidos para tratar dados geográficos!

SGBD Relacional

- Modelo de Dados Relacional
  - Banco de dados é organizado em uma coleção de relações ou tabelas relacionadas entre si.

Aluno

MATRICULA	NOME	CURSOID
98765	João	MAT
67765	José	BIO
84562	Maria	ENG
34256	Luis	INFO
3452672	Ana	MAT
34529	Luana	MAT

Curso

CURSOID	TITULO	DURAÇÃO
INFO	Informática Indust.	4
BIO	Biologia	4
ENG	Engenharia Civil	5
MAT	Licenciatura Mat.	4

SGBD Objeto-Relacional

- Modelo de Dados Objeto-Relacional:
  - É uma extensão do Modelo Relacional com conceitos da modelagem Orientada por Objeto.
  - Combina os benefícios dos dois modelos.
- Fornecem suporte para:
  - Criar objetos complexos
  - Executar consultas avançadas sobre dados complexos
- A linguagem de consulta OR é uma extensão da linguagem SQL para suportar o modelo de objetos

SGBD Objeto-Relacional

Exemplo: Oracle

```
CREATE TYPE SDO_GEOMETRY AS OBJECT (
SDO_GTYPE       NUMBER,
SDO_SRID        NUMBER,
SDO_POINT       SDO_POINT_TYPE,
SDO_ELEM_INFO   SDO_ELEM_INFO_ARRAY,
SDO_ORDINATES   SDO_ORDINATE_ARRAY);
```

```
CREATE TABLE Estados (
SIGLA    VARCHAR(2)
POP      NUMBER(10,10)
GEOM     SDO_GEOMETRY)
```

Estados

SIGLA	POP	GEOM
MG	222222	
RJ	333333	
SP	444444	

Extensão Espacial

- SGBD-OR são estendidos para suportar:
  - Tipos de dados espaciais: polígono, ponto, linha, raster, etc;
  - Operadores e funções utilizados na SQL para manipular dados espaciais (consultas e junção)
  - Métodos eficientes de acesso aos dados espaciais

- Extensões existentes (seguem padrão OGC):
  - Comerciais
    - Oracle Spatial
    - IBM DB2 Spatial Extender
  - Livres
    - PostGIS
    - Extensão espacial para MySQL

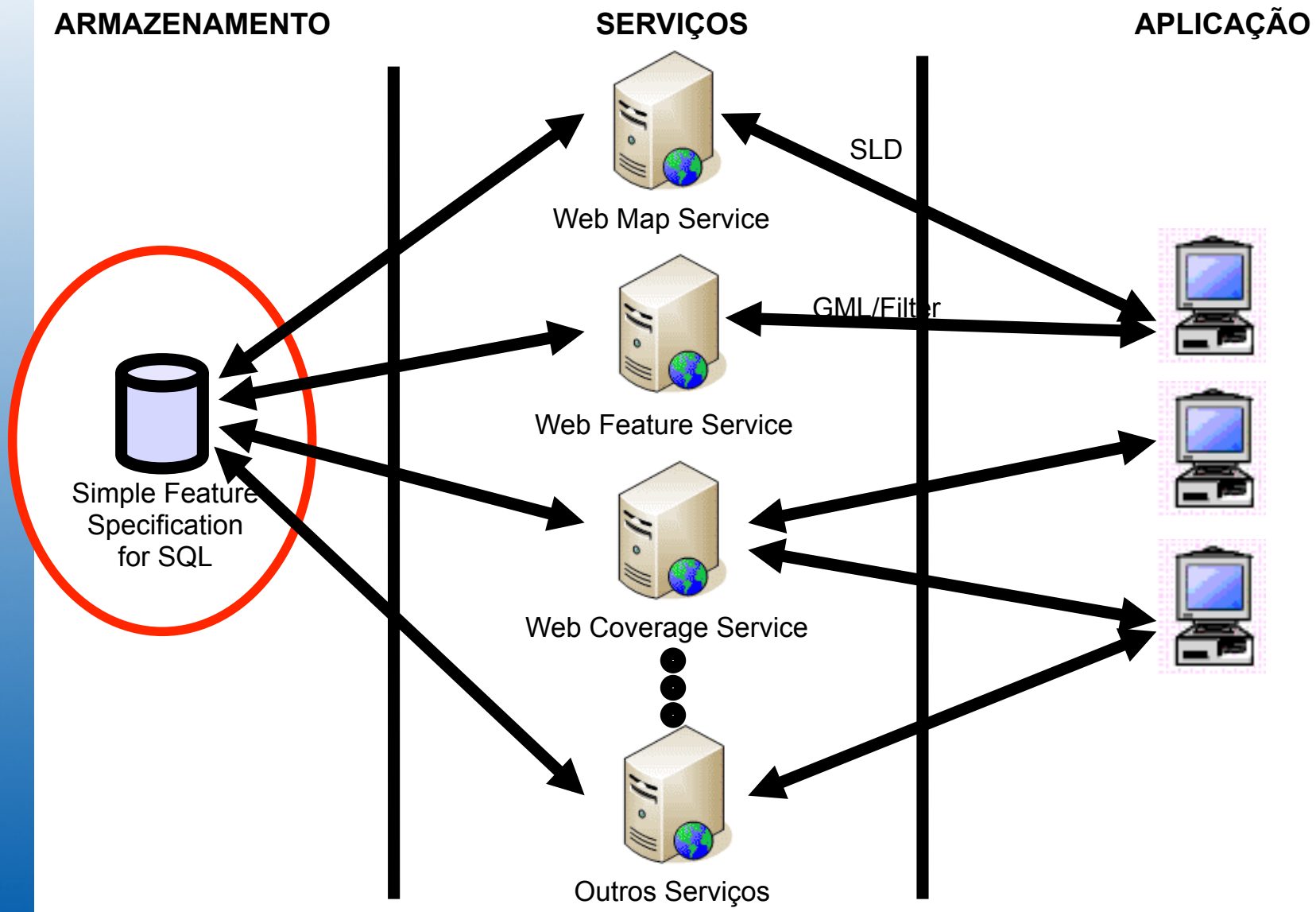
Open Geospatial Consortium – OGC

- Consórcio formado por empresas, universidades e agências governamentais.
- Promover o desenvolvimento de padrões que facilitem a interoperabilidade entre sistemas envolvendo informação geo-espacial.
- Os produtos do trabalho do OGC são apresentados sob a forma de especificações de interfaces e padrões de intercâmbio.

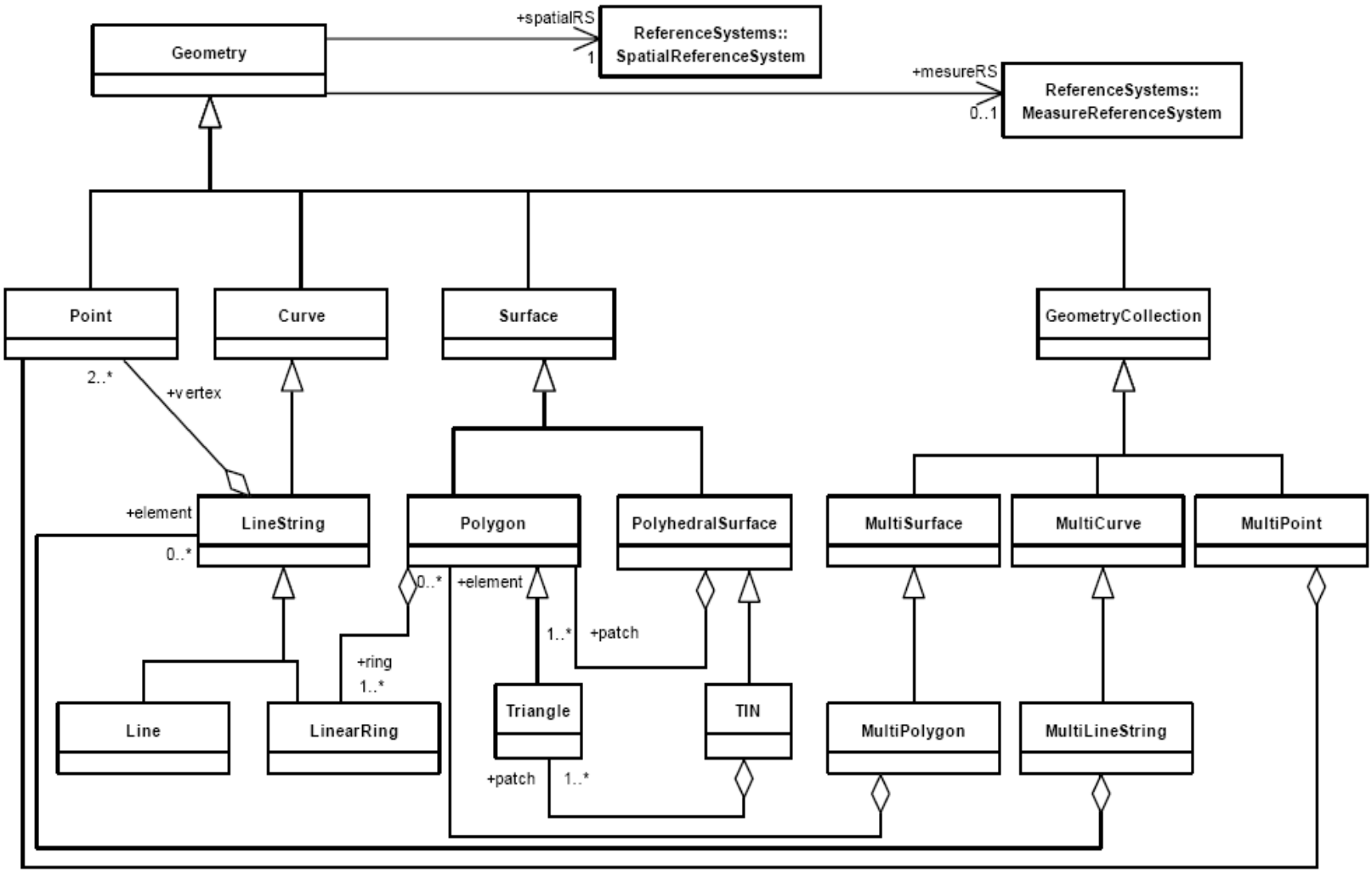
Open Geospatial Consortium – OGC

- Algumas especificações OGC:
  - **SFS-SQL** (Simple Feature Specification For SQL): especificações sobre o armazenamento e recuperação de dados espaciais vetoriais em sistemas de bancos de dados.
  - **GML** (Geography Markup Language): intercâmbio de dados.
  - **OWS** (OGC Web Services): especificações de serviços WEB
    - WFS: Web Feature Service
    - WMS: Web Map Server

OGC: Panorama Geral



SFS: Modelo Geométrico



Fonte da Figura: OGC

SFS: Modelo Geométrico

- Tipos de geometrias vetoriais - exemplos:
  - Criar uma tabela para armazenar os municípios de São Paulo:

```
CREATE TABLE municipiossp
(cod           INTEGER,
nomemunicp     VARCHAR(255) NULL,
populacao      REAL
geometria      POLYGON);
```

  - Criar uma tabela para armazenar os rios de São Paulo:

```
CREATE TABLE drenagemsp
( cod          INTEGER,
nomerio        VARCHAR(255) NULL,
geometria      LINESTRING);
```

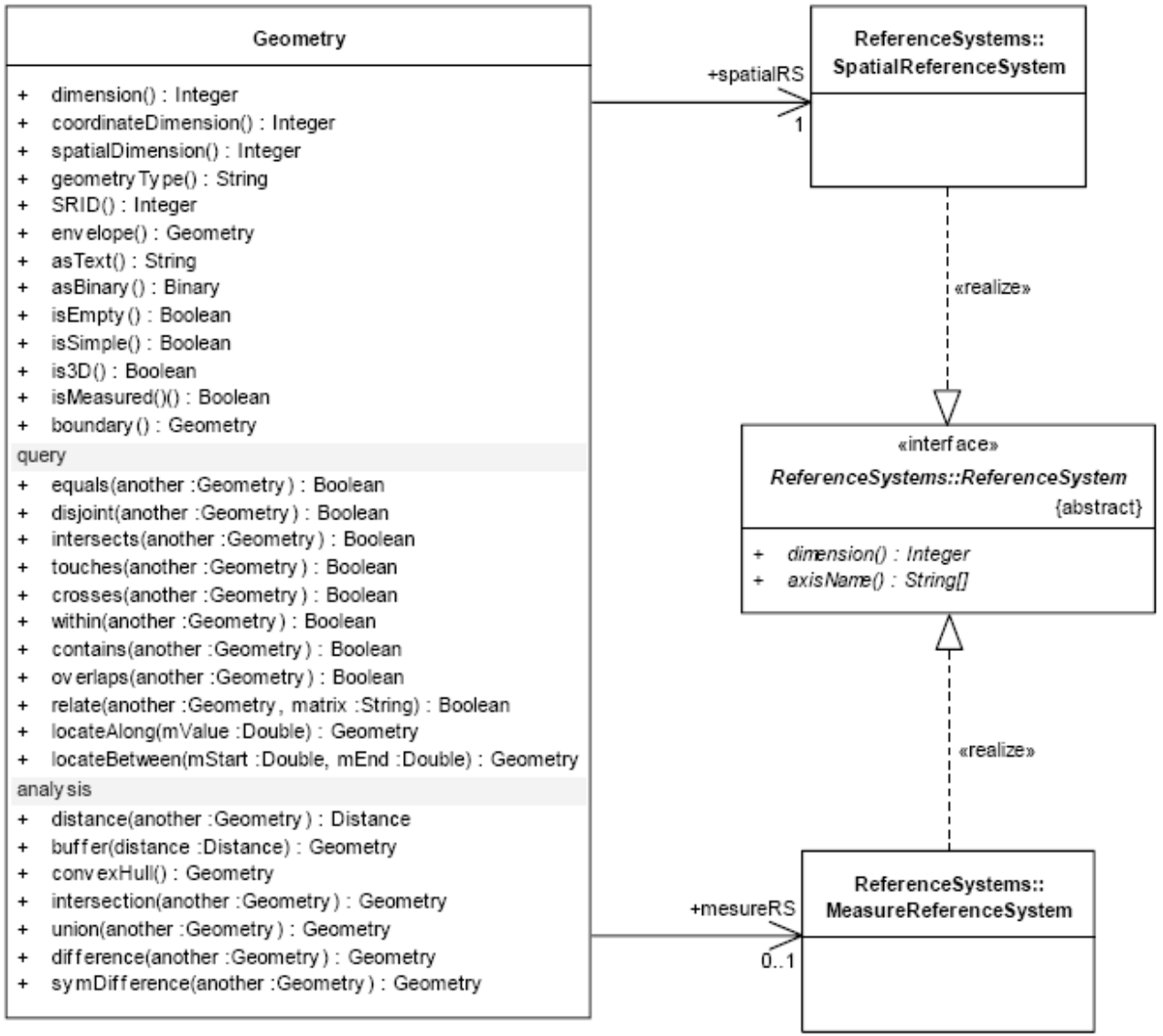
SFS: Modelo Geométrico

- Dimensionalidade da figura geométrica:
  - 0-dimensional: ponto
  - 1-dimensional: linha
  - 2-dimensional: polígono

  Obs: não diz respeito ao espaço onde a figura está inserida!

- Dimensionalidade do espaço onde está inserida a figura geométrica:
  - 2D: plano cartesiano
  - 3D: espaço tridimensional
  - 2D+M: plano cartesiano com uma medida associada
  - 3D+M: espaço tridimensional com uma medida associada

SFS: Geometria



Fonte da Figura: OGC

SFS: Ponto e Linha

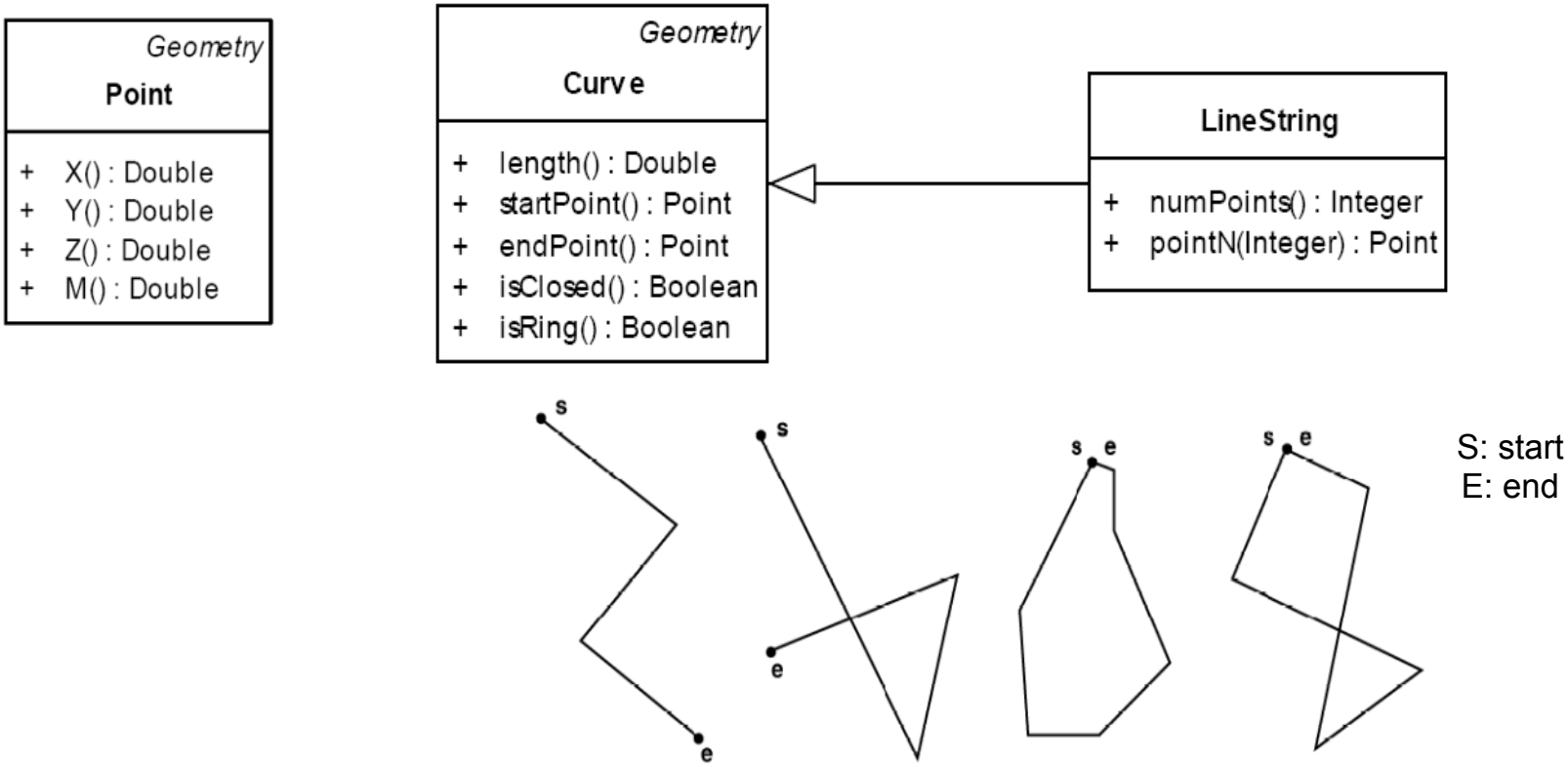
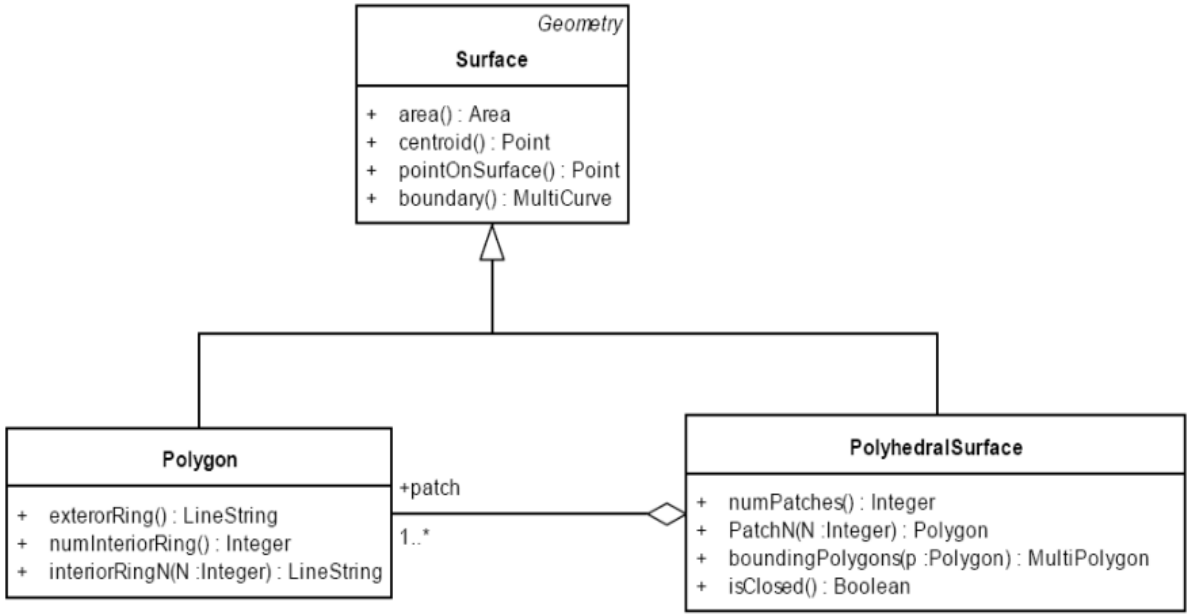
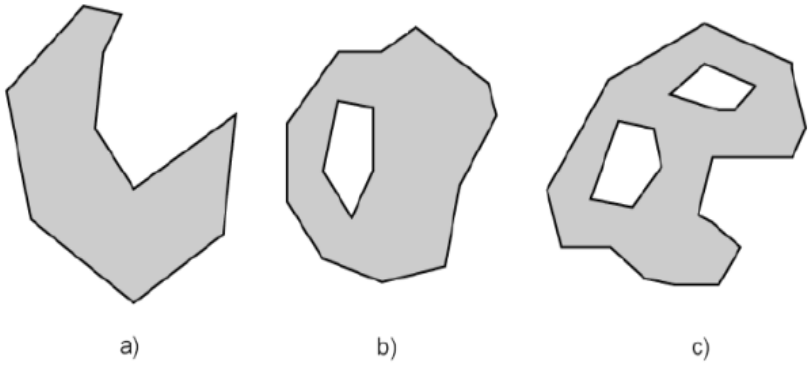


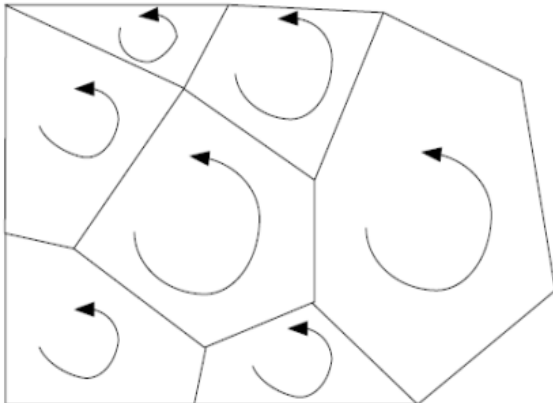
Figure 6: Examples of LineStrings
Simple LineString (a),
Non-simple LineString (b),
Simple, closed LineString (a LinearRing) (c),
Non-simple closed LineString (d)

Fonte da Figura: OGC

SFS: Polígono







Fonte da Figura: OGC

SFS: Polígono

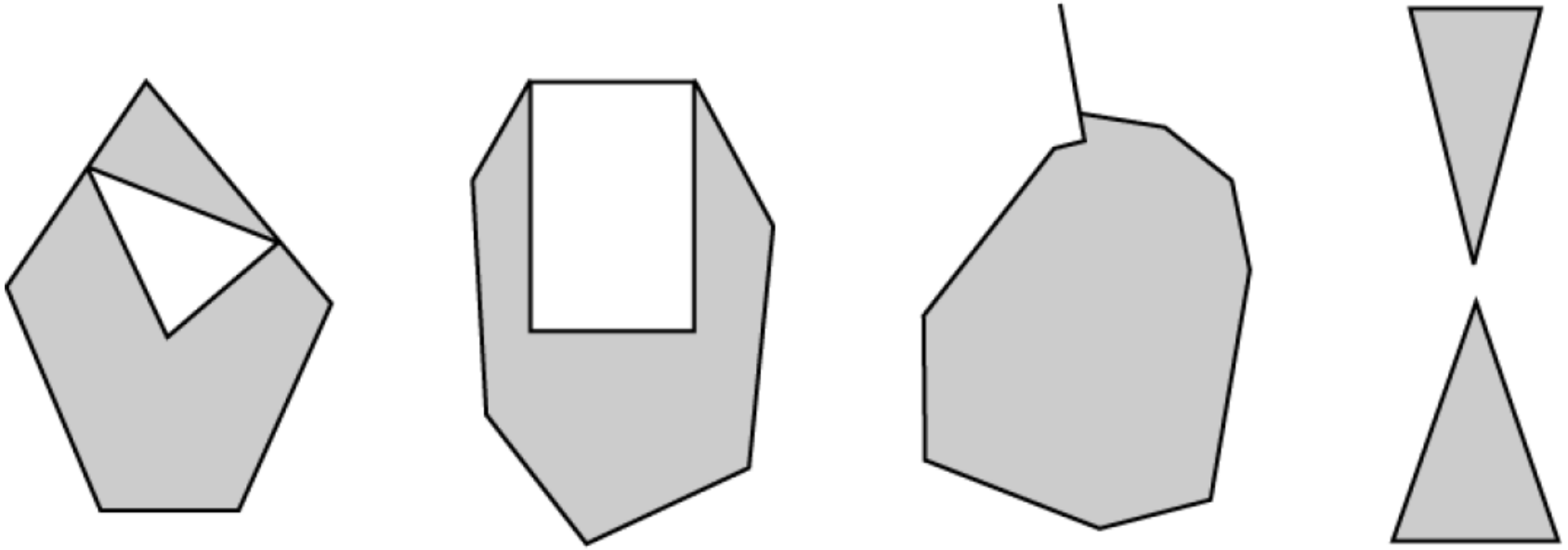
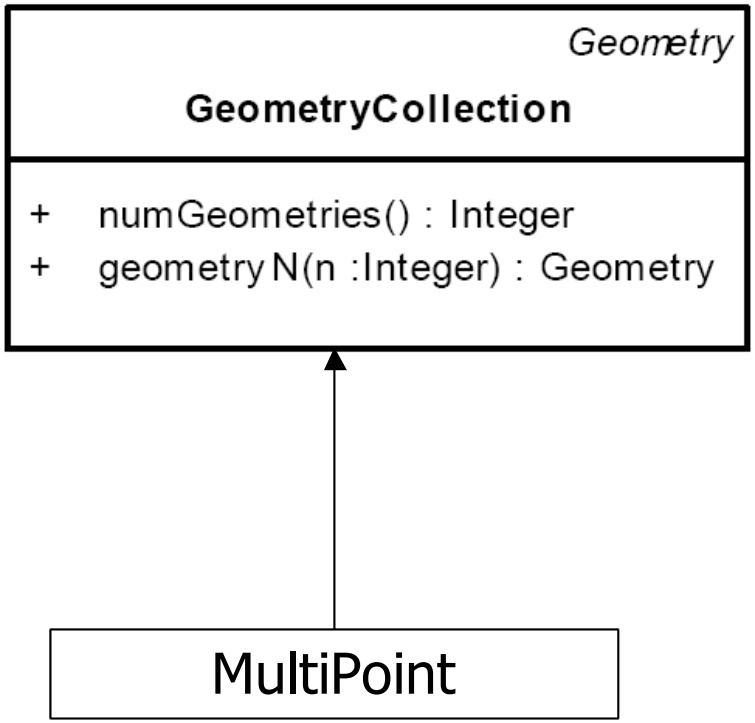
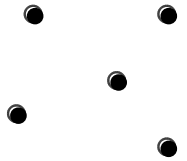


Figure 12: Examples of objects not representable as a single instance of Polygon

Fonte da Figura: OGC

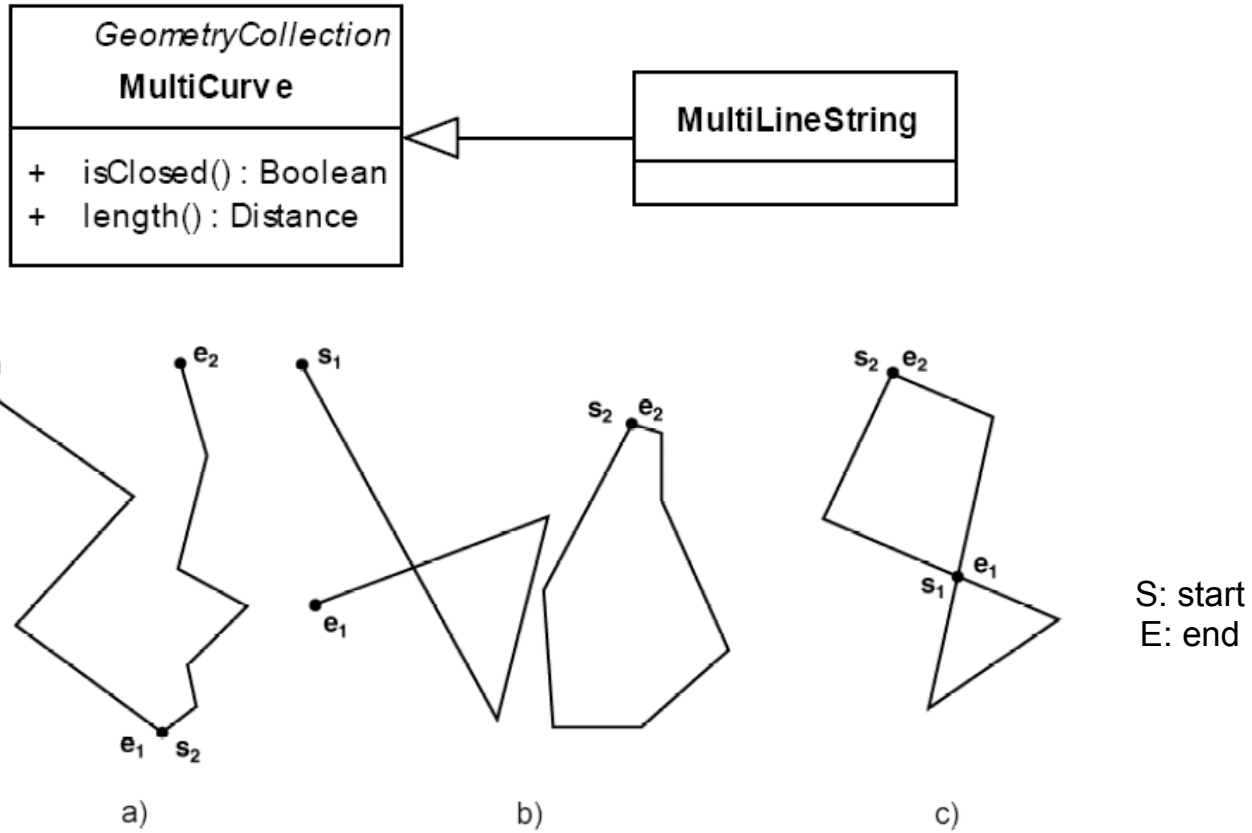
SFS: Coleções Homogêneas





Fonte da Figura: OGC

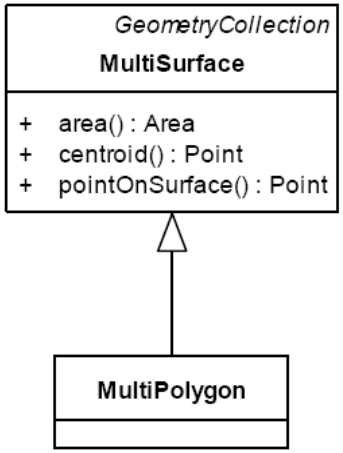
SFS: Coleções Homogêneas



(a) Simple MultiLineString,
(b) Non-simple MultiLineString with 2 elements,
(c) Non-simple, closed MultiLineString with 2 elements

Fonte da Figura: OGC

SFS: Coleções Homogêneas



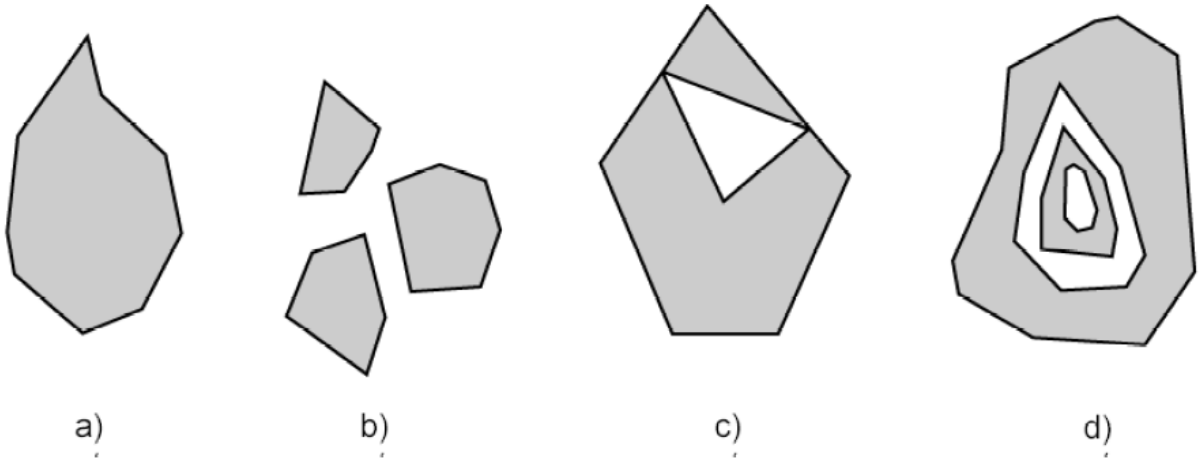


Figure 17: Examples of MultiPolygons with 1 (a), 3 (b) , 2 (c) and 2 (d) Polygon elements

Fonte da Figura: OGC

SFS: Coleções Homogêneas

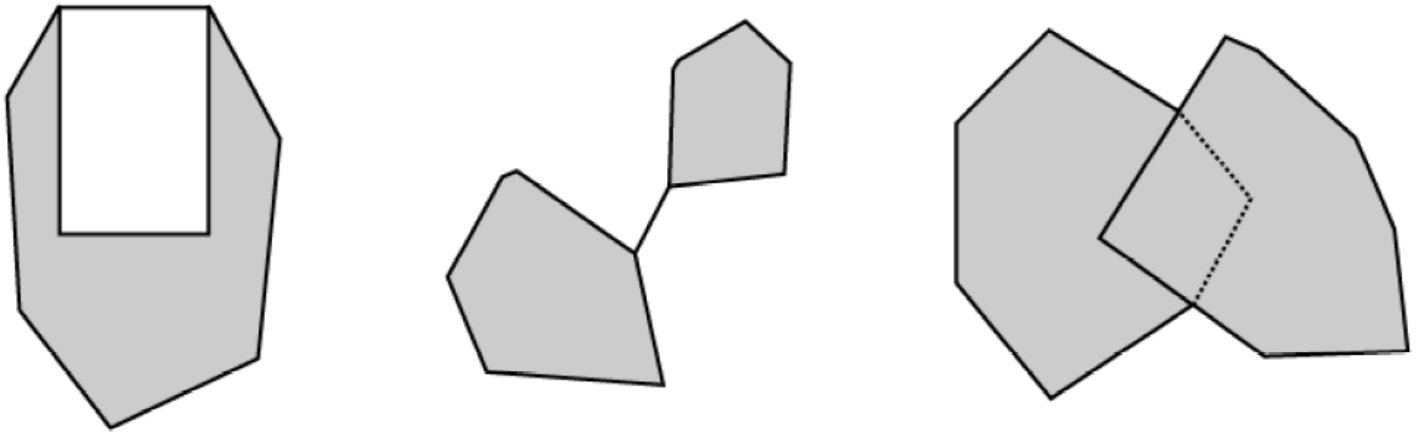


Figure 18: Geometric objects not representable as a single instance of a MultiPolygon

Fonte da Figura: OGC

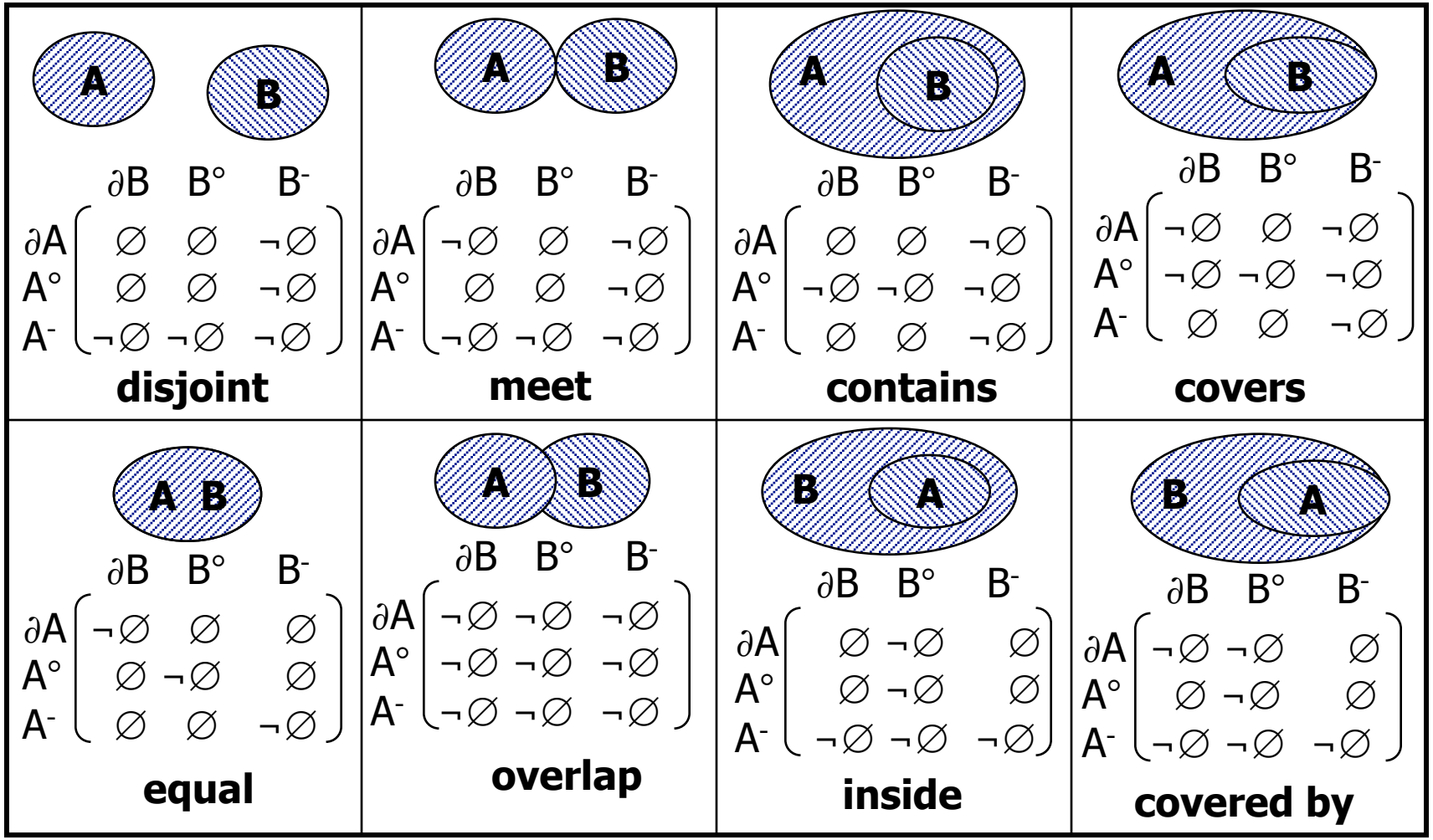
SFS: Coleções Heterogêneas

- São permitidas coleções heterogêneas como instâncias da classe GeometryCollection

SFS: Operadores Topológicos

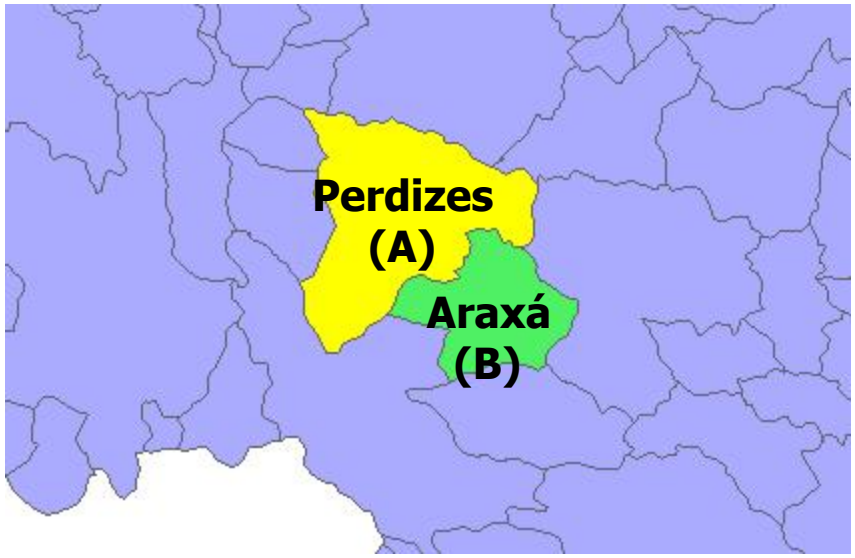
- Operadores topológicos baseados na matriz de 9-Interseções estendida dimensionalmente (DE-9IM) : `touches`, `equals`, `overlaps`, `disjoints`, `intersects`, `contains`, `insides`, `covers`, `coveredBy`.

Predicados topológicos: Matriz 9-Interseções



Fonte: Adaptado de Egenhofer e Herring (1991)

Predicados topológicos: Matriz 9-Interseções



$$
\begin{array}{c} \\ \partial A \\ A^{\circ} \\ A^{-} \end{array}
\begin{array}{c}
\begin{array}{ccc} \partial B & B^{\circ} & B^{-} \end{array} \\
\begin{pmatrix} \neg\varnothing & \varnothing & \neg\varnothing \\ \varnothing & \varnothing & \neg\varnothing \\ \neg\varnothing & \neg\varnothing & \neg\varnothing \end{pmatrix}
\end{array}
$$

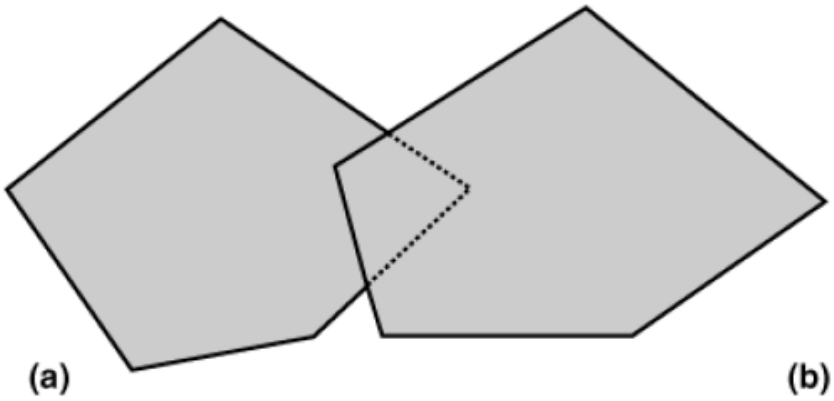
TOCA

Fonte: Gilberto Ribeiro (2006)

DE-9IM - Matriz de 9-Interseções estendida dimensionalmente

Table 1: The DE-9IM

	Interior	Boundary	Exterior
Interior	$dim(I(a) \cap I(b))$	$dim(I(a) \cap B(b))$	$dim(I(a) \cap E(b))$
Boundary	$dim(B(a) \cap I(b))$	$dim(B(a) \cap B(b))$	$dim(B(a) \cap E(b))$
Exterior	$dim(E(a) \cap I(b))$	$dim(E(a) \cap B(b))$	$dim(E(a) \cap E(b))$



	Interior	Boundary	Exterior
Interior	2	1	2
Boundary	1	0	1
Exterior	2	1	2

Dim = -1 => vazio
Dim = 0 => ponto
Dim = 1 => linha
Dim = 2 => polígono.

Fonte da Figura: OGC

DE-9IM - Matriz de 9-Interseções estendida dimensionalmente

The pattern matrix consists of a set of nine pattern-values, one for each cell in the matrix. The possible pattern values of *p* are {T, F, *, 0, 1, 2} and their meanings for any cell where x is the intersection set for the cell are as follows:

$$
\begin{aligned}
&p = T \Rightarrow \dim(x) \in \{0, 1, 2\}, \text{ i.e. } x \neq \varnothing \\
&p = F \Rightarrow \dim(x) = -1, \text{ i.e. } x = \varnothing \\
&p = * \Rightarrow \dim(x) \in \{-1, 0, 1, 2\}, \text{ i.e. Don't Care} \\
&p = 0 \Rightarrow \dim(x) = 0 \\
&p = 1 \Rightarrow \dim(x) = 1 \\
&p = 2 \Rightarrow \dim(x) = 2
\end{aligned}
$$

Fonte da Figura: OGC

DE-9IM - Touches

a.Touch(b)⇔ a.Relate(b, “FT*******”)
∨
a.Relate(b, “F**T*****”) ∨
a.Relate(b, “F***T****”)

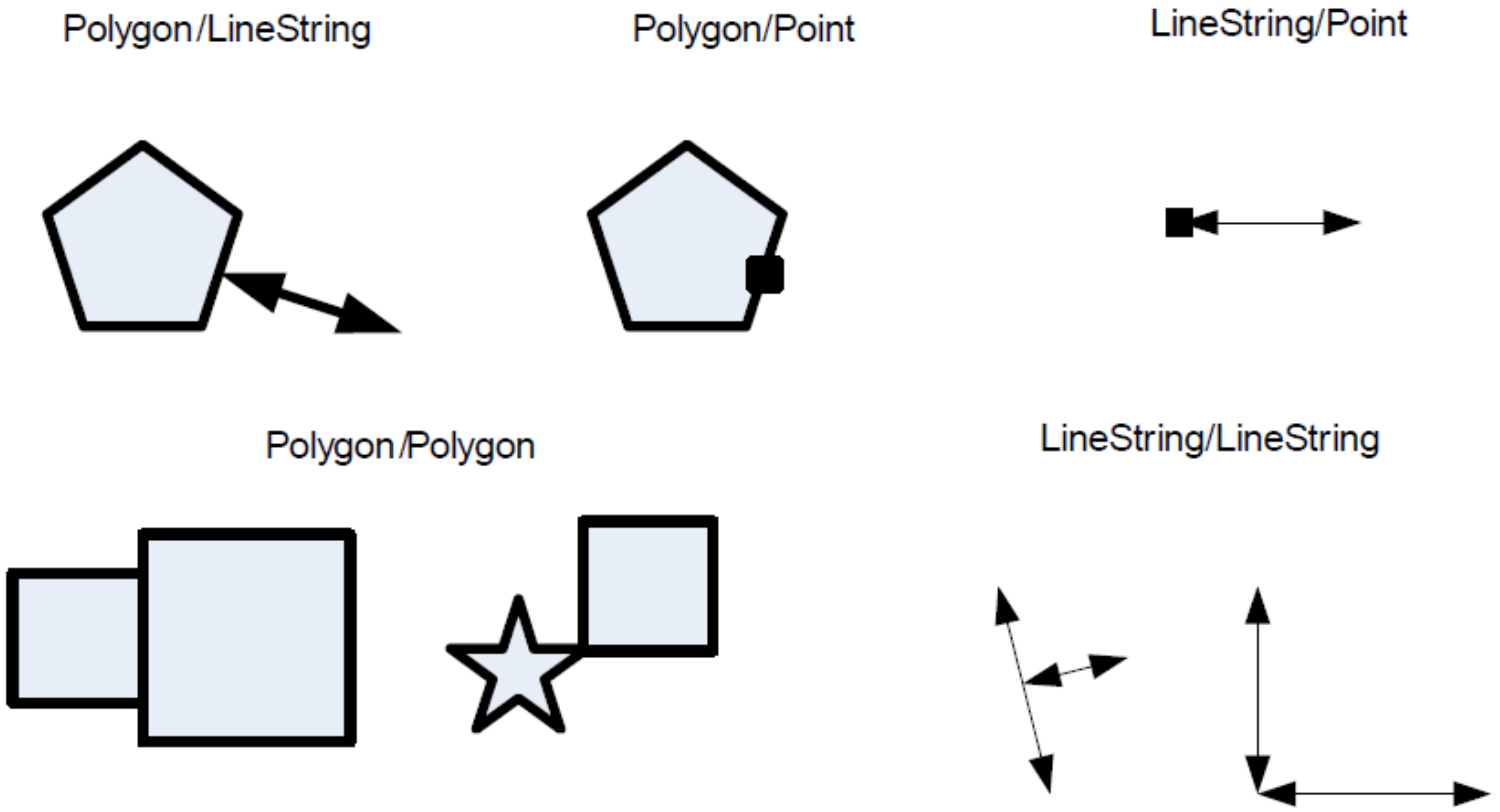


Figure 20: Examples of the Touches relationship

Fonte da Figura: OGC

DE-9IM - Within

a.Within(b) ⇔ a.Relate(b, “T*F**F***”)

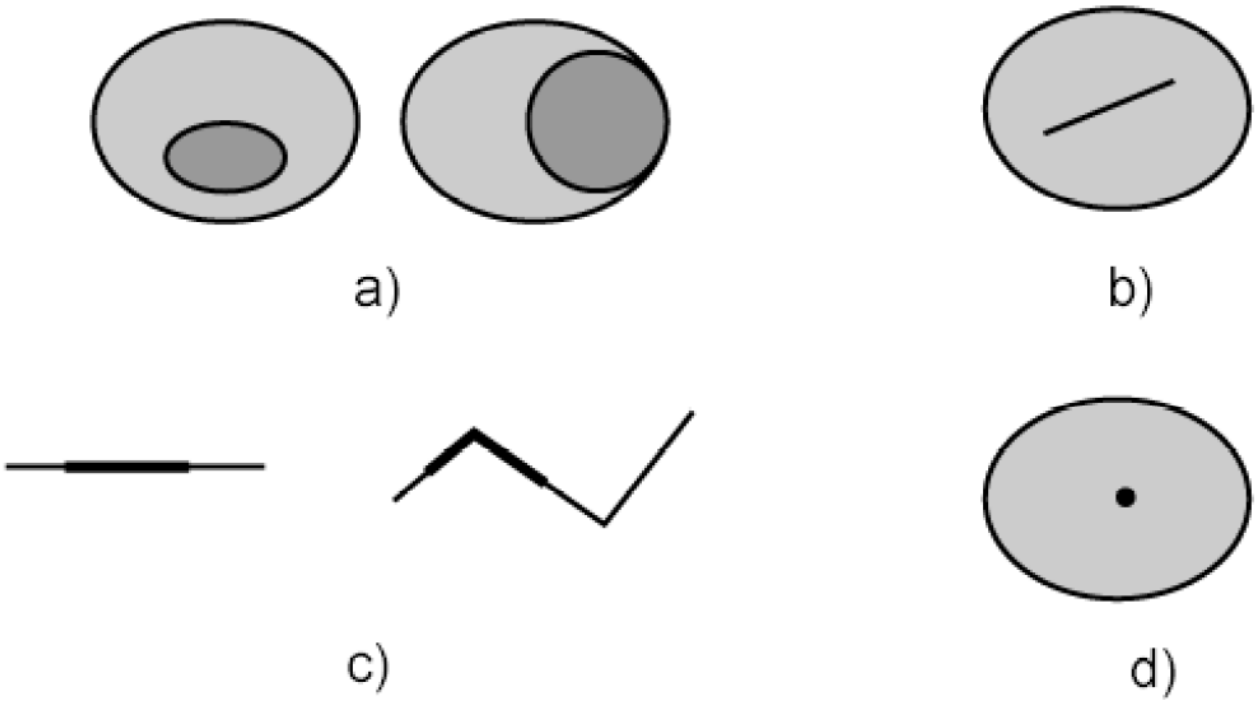


Figure 22: Examples of the “Within” relationship
Polygon/Polygon (a), Polygon/LineString (b), LineString/LineString (c), and Polygon/Point (d)

Fonte da Figura: OGC

DE-9IM - Overlaps

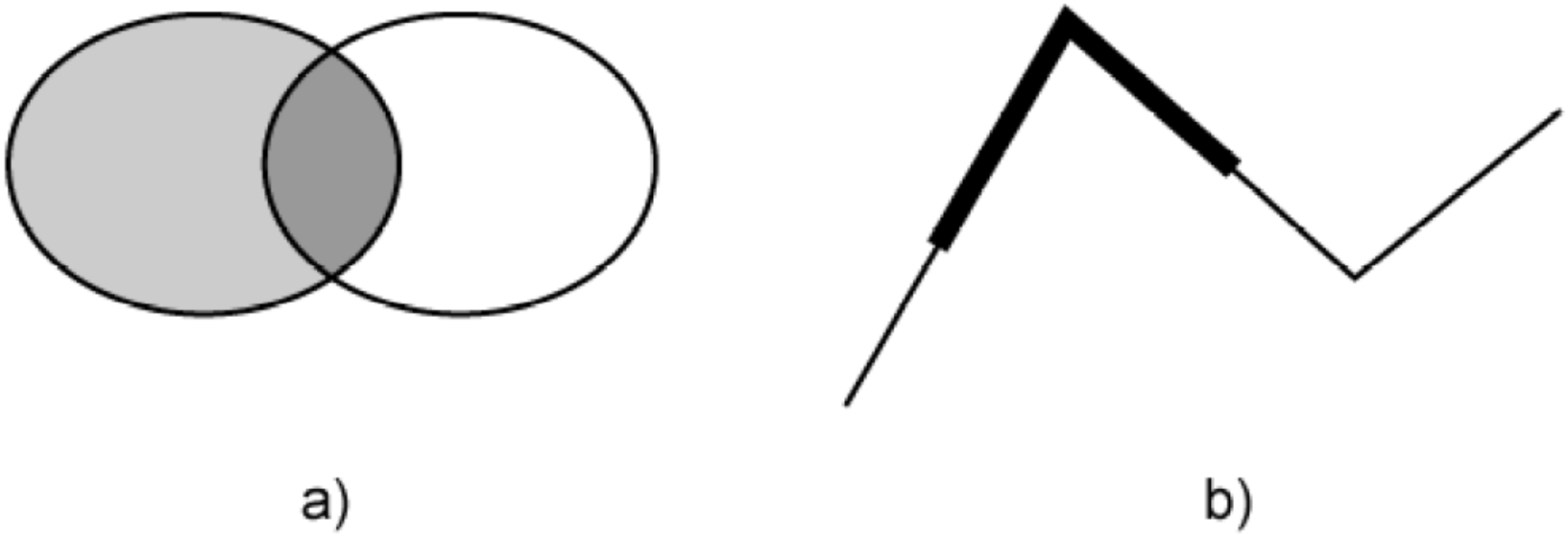


Figure 23: Examples of the Overlaps relationship Polygon/LineString (a) and LineString/LineString (b)

Fonte da Figura: OGC

DE-9IM - Crosses

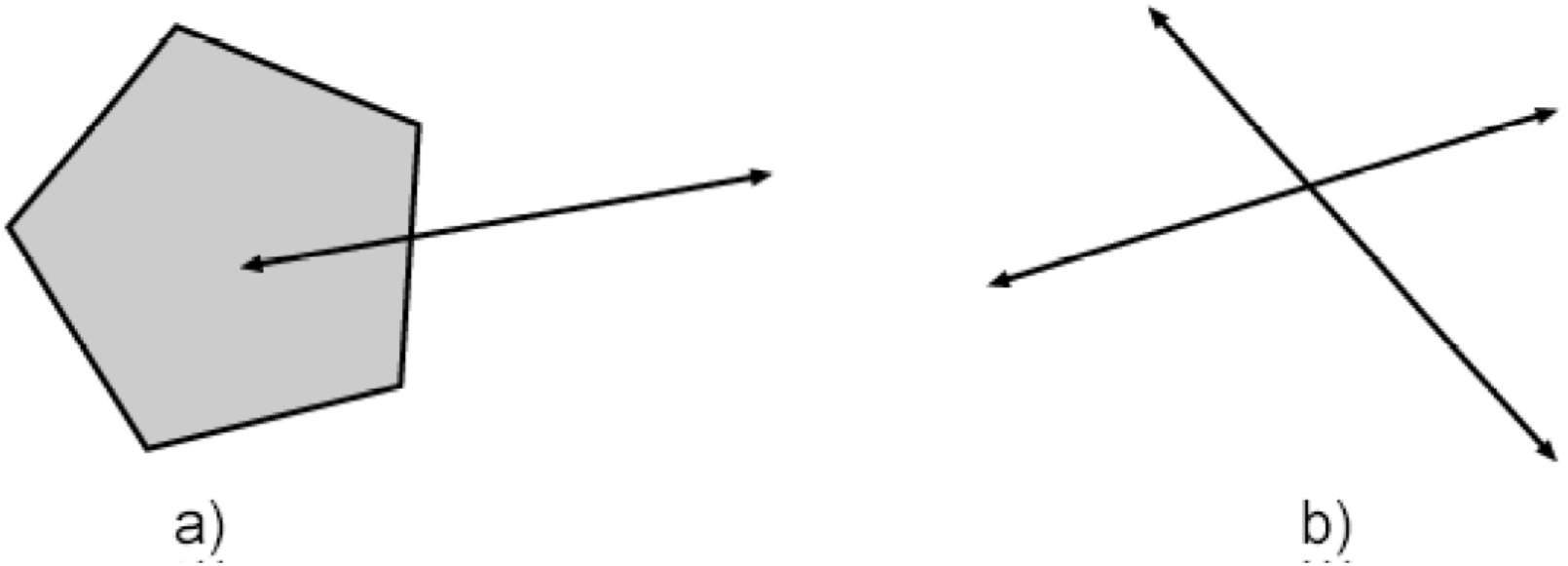


Figure 21: Examples of the Crosses relationship Polygon/LineString (a) and LineString/LineString (b)

Fonte da Figura: OGC

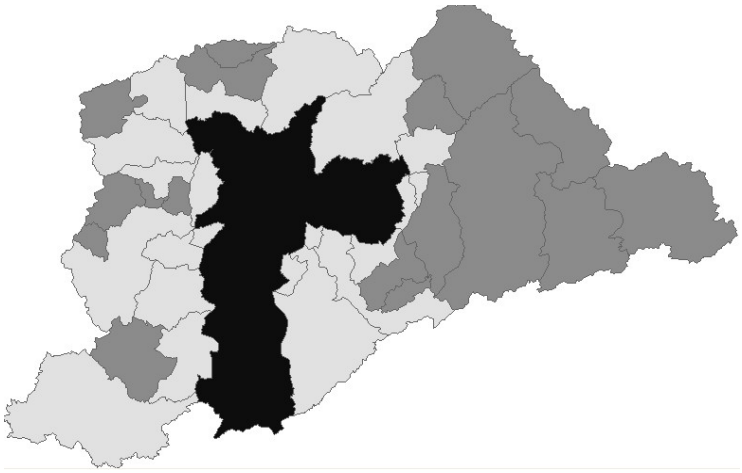
SFS: Outros Operadores

- Outros operadores: `distance`, `buffer`, `convexHull`, `intersection`, `union`, `difference`, `area`, `centroid` e `pointOnSurface` ....

SFS-SQL

- Spatial SQL - Exemplos:
  - Encontre todos os municípios de SP que são vizinhos do município de São Paulo:

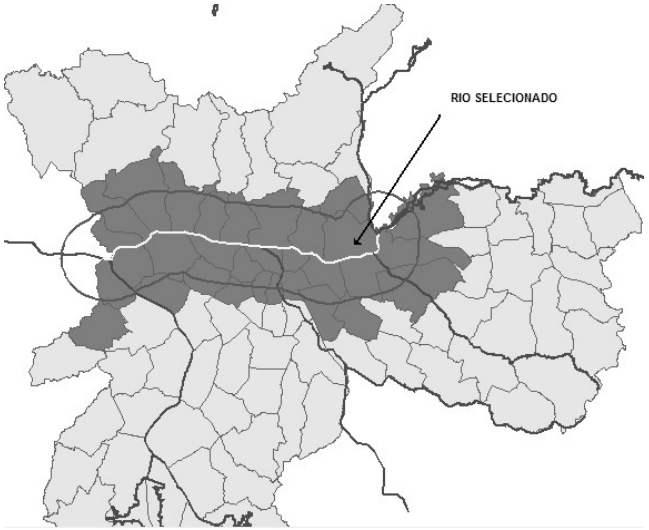
```
SELECT d2.nomemunicp
FROM municipiossp d1, municipiossp d2
WHERE TOUCHES(d1.geometria, d2.geometria)
AND (d2.nomemunicp <> 'SAO PAULO')
AND (d1.nomemunicp = 'SAO PAULO');
```



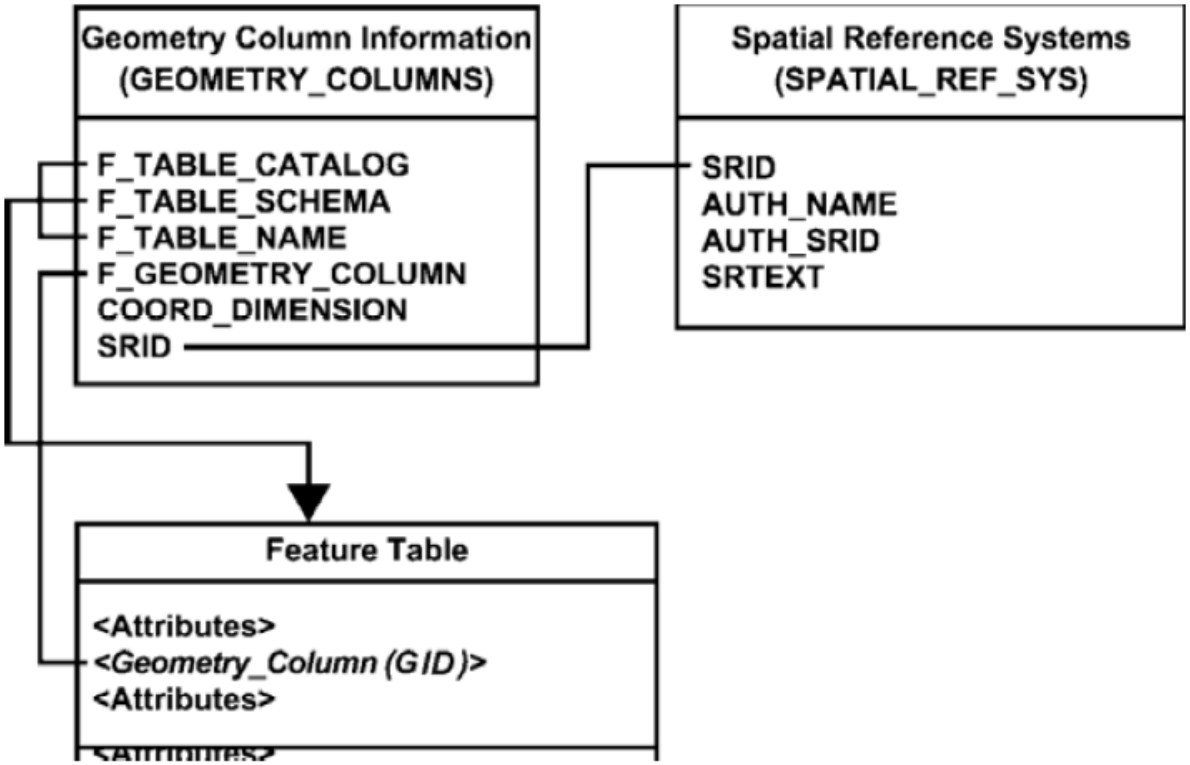
SFS-SQL

- Spatial SQL - Exemplos:
  - Encontre todos os municípios de SP que estão num raio de 3Km do rio X:

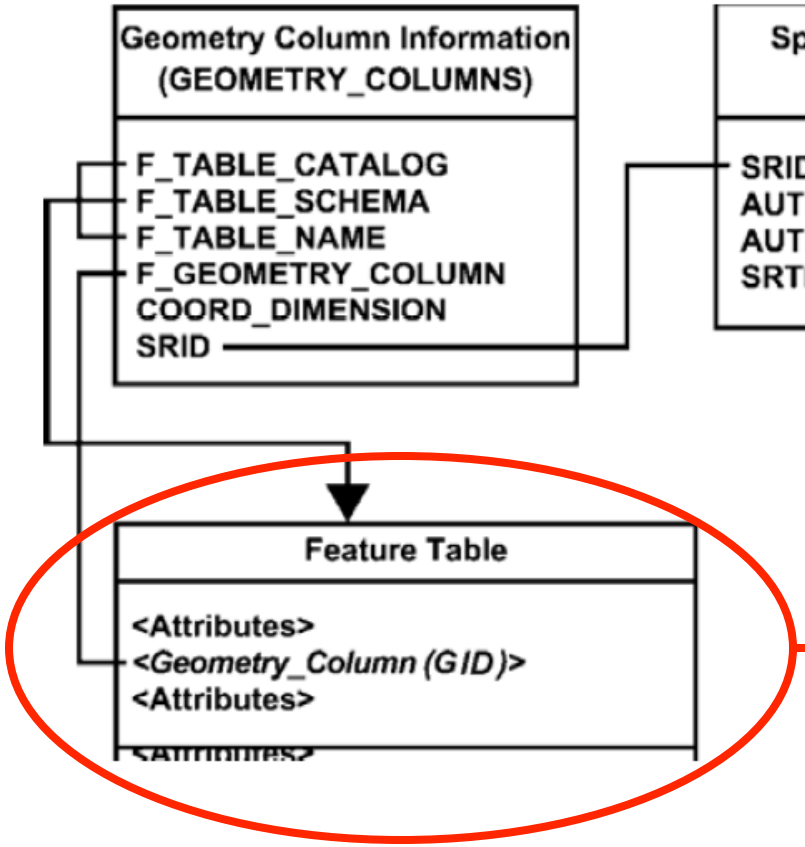
```
SELECT nomemunicp
FROM          municipiossp, drenagemsp
WHERE  INTERSECTS
(BUFFER(drenagemsp.geometria, 3000),
municipiossp.geometria)
AND    drenagemsp.nomerio = 'X';
```



SFS-SQL – Metadata tables

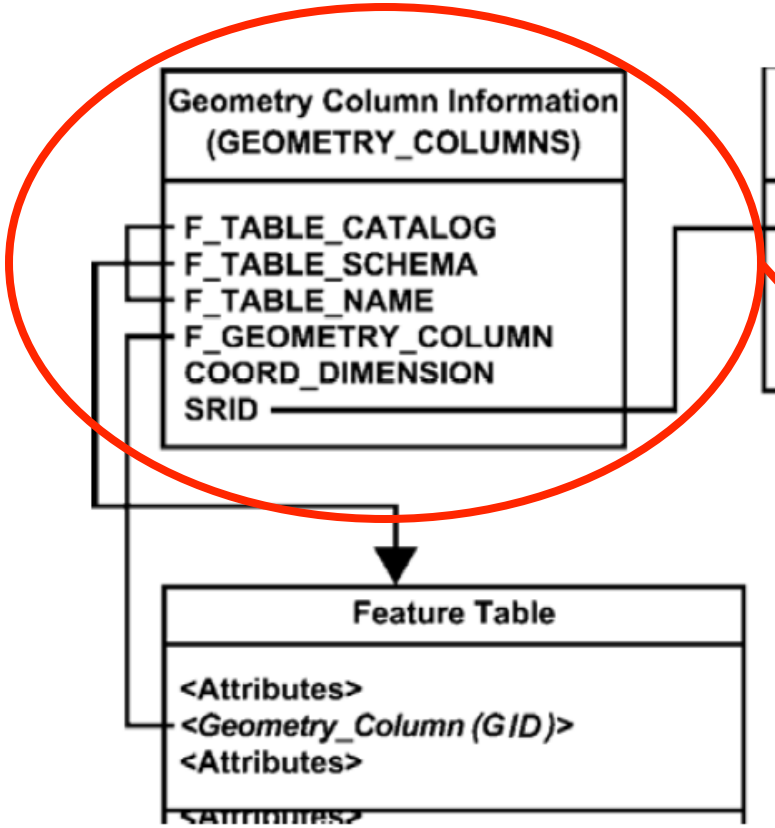


SFS-SQL – Metadata tables



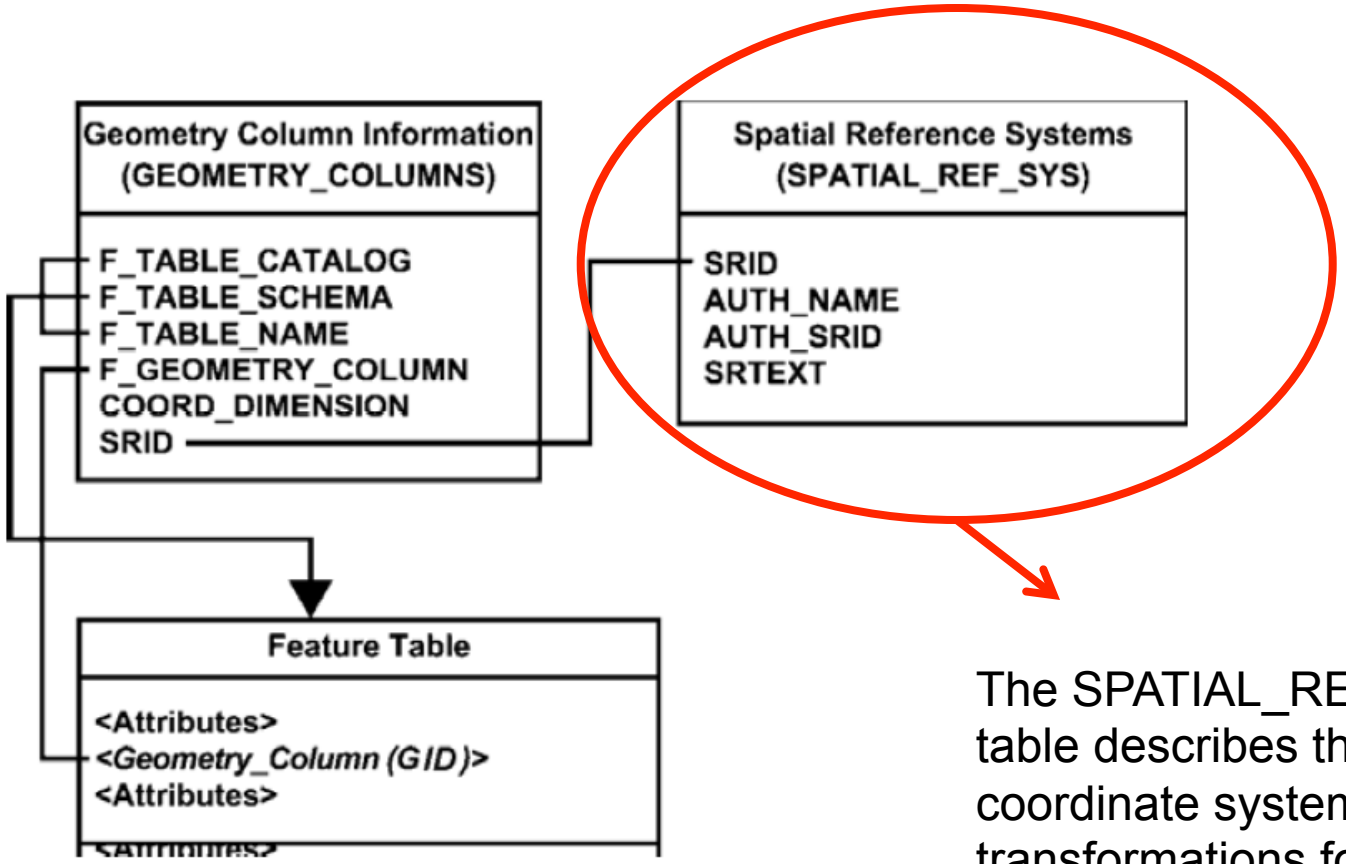
The feature table stores a collection of features. A feature table's columns represent feature attributes, while rows represent individual features. The Geometry of a feature is one of the feature attributes, and is an SQL Geometry Type.

SFS-SQL – Metadata tables



The GEOMETRY_COLUMNS table describes the available feature tables and their Geometry properties.

SFS-SQL – Metadata tables



The SPATIAL_REF_SYS table describes the coordinate system and transformations for Geometry

WKT – Well-known Text Representation

Each Geometry Type has a Well-known Text Representation that can be used both to construct new instances of the type and to convert existing instances to textual form for alphanumeric display.

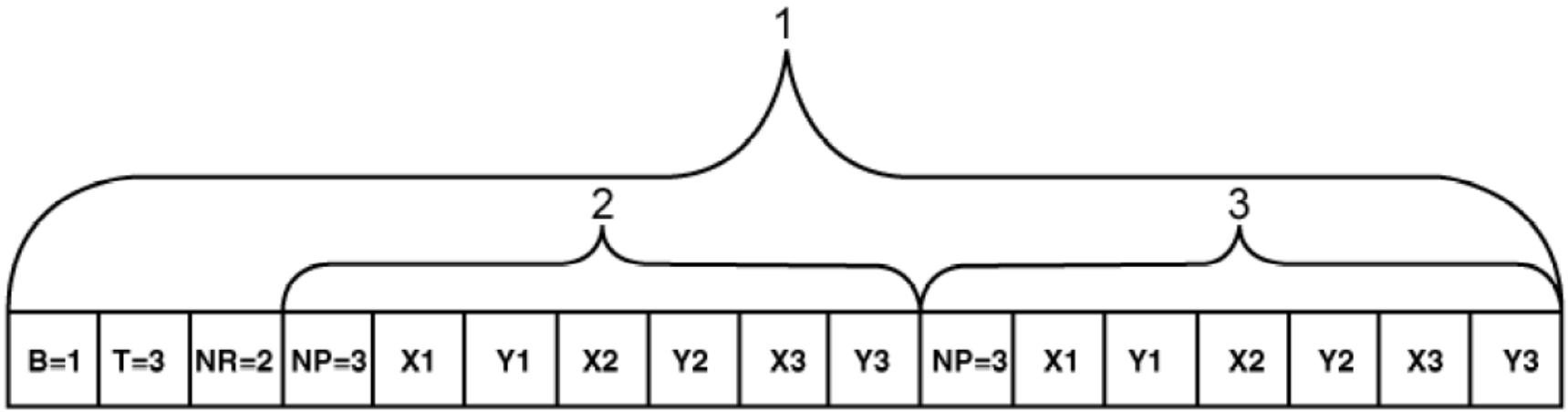
WKT – Well-known Text Representation

Geometry Type	Text Literal Representation	Comment
Point	Point (10 10)	a Point
LineString	LineString (10 10, 20 20, 30 40)	a LineString with 3 points
Polygon	Polygon ((10 10, 10 20, 20 20, 20 15, 10 10))	a Polygon with 1 exteriorRing and 0 interiorRings
Multipoint	MultiPoint ((10 10), (20 20))	a MultiPoint with 2 points
MultiLineString	MultiLineString ((10 10, 20 20), (15 15, 30 15))	a MultiLineString with 2 linestrings
MultiPolygon	MultiPolygon (((10 10, 10 20, 20 20, 20 15, 10 10)), ((60 60, 70 70, 80 60, 60 60)))	a MultiPolygon with 2 polygons

WKB – Well-known Binary Representation

The Well-known Binary Representation for Geometry (WKBGeometry) provides a portable representation of a geometric object as a contiguous stream of bytes. It permits geometric object to be exchanged between an SQL/CLI client and an SQL-implementation in binary form.

WKB – Well-known Binary Representation



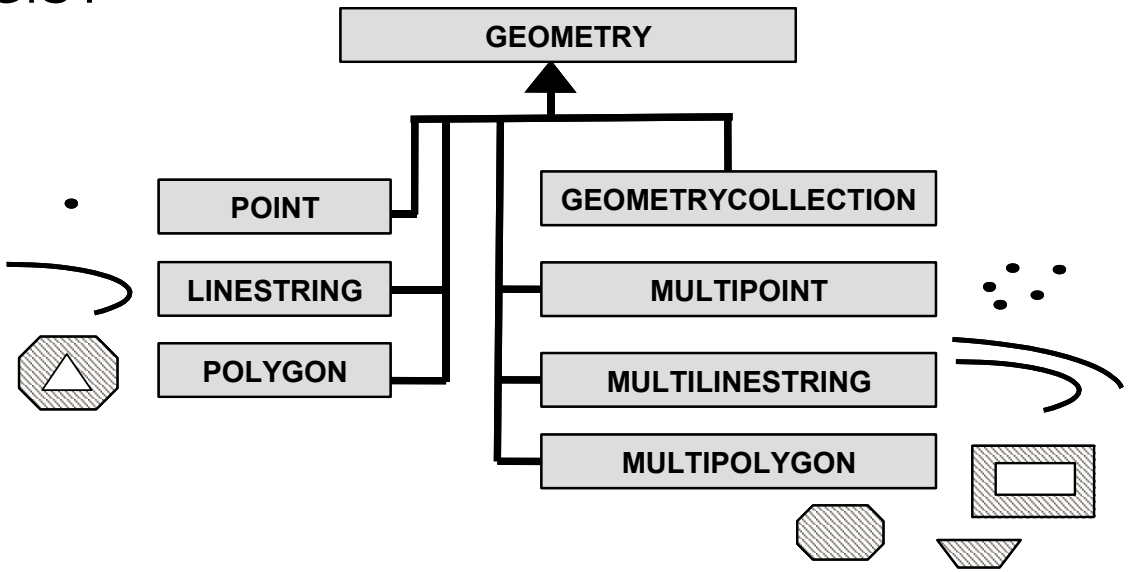
Key

1 WKB Polygon
2 ring 1
3 ring 2

Figure 25: Well-known Binary Representation for a geometric object in NDR format (B = 1) of type Polygon (T = 3) with 2 LinearRings (NR = 2) each LinearRing having 3 points (NP = 3)

Extensão espacial - PostGIS

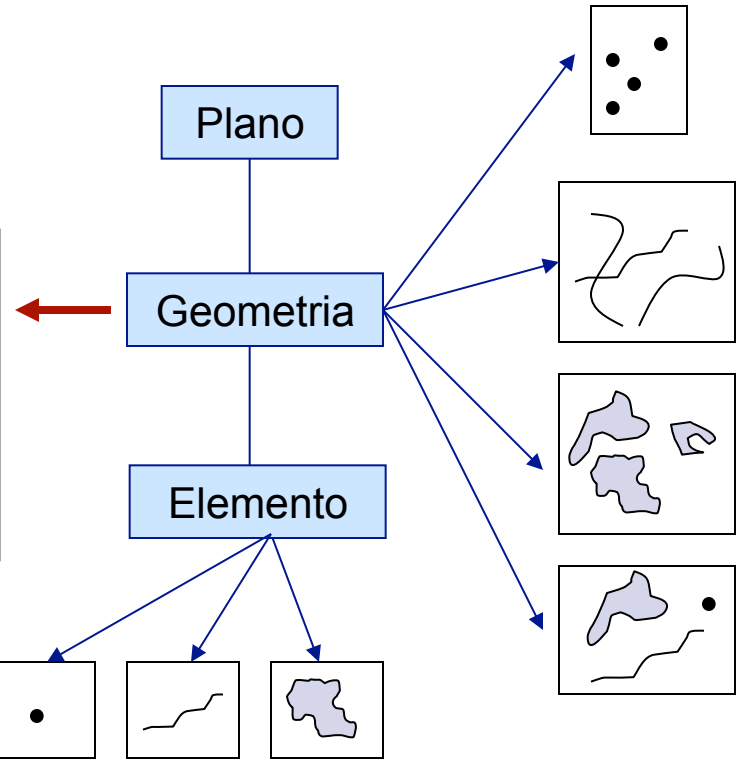
- Extensão do SGBD PostgreSQL (SFS-SQL):
  - Tipos de dados geométricos.
  - Operadores espaciais:
    - Através da biblioteca GEOS
  - Métodos de Acesso Espacial:
    - R-Tree sobre GiST



Extensão espacial - Oracle Spatial

- Extensão do SGBD Oracle (SFS-SQL):
  - Tipos de dados geométricos.
  - Operadores e funções espaciais:
  - Métodos de Acesso Espacial:
    - R-Tree e QuadTree

```
CREATE TYPE SDO_GEOMETRY AS OBJECT (
SDO_GTYPE          NUMBER,
SDO_SRID           NUMBER,
SDO_POINT          SDO_POINT_TYPE,
SDO_ELEM_INFO      SDO_ELEM_INFO_ARRAY,
SDO_ORDINATES      SDO_ORDINATE_ARRAY);
```



Fonte: Karine Ferreira (2006)