PTR 5003
Informações Espaciais

Aula SIG:
Conceitos Básicos,
Modelos de Representação e
Coleta de Dados

1quadr19

Sugestão de Bibliografia

CÂMARA, G; DAVIS, C.; MONTEIRO, A.M. (ed). **Introdução à Ciência da Geoinformação**. São José dos Campos, INPE, 2004. (Disponível online: http://www.dpi.inpe.br/gilberto/livro/introd/index.html)

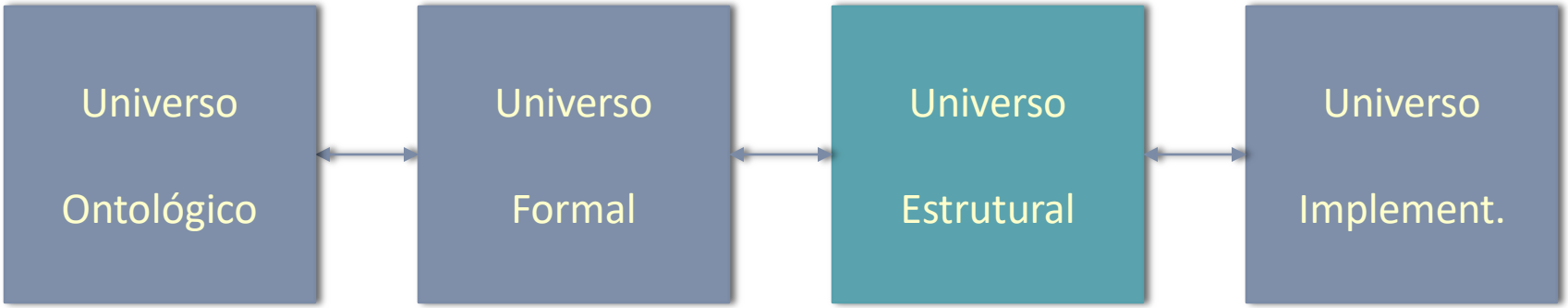
CASANOVA, Marco (Org.) ; CÂMARA, G. (Org.) ; DAVIS, Clodoveu (Org.) ; VINHAS, Lúbia (Org.) ; QUEIROZ, Gilberto Ribeiro (Org.) . **Bancos de Dados Geograficos**. 1. ed. Curitiba: Editora Mundo Geo, 2005. v. 1. 504 p. (Disponível online: http://www.dpi.inpe.br/gilberto/livro/bdados/index.html)

LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2013, 539p.

BURROUGH, P. A.; MCDONNELL, R. A. **Principles of geographical information systems**. New York: Oxford University Press, 1998. 334p.

Caderno Didático do INPE sobre SIG

Processo de Representação Computacional

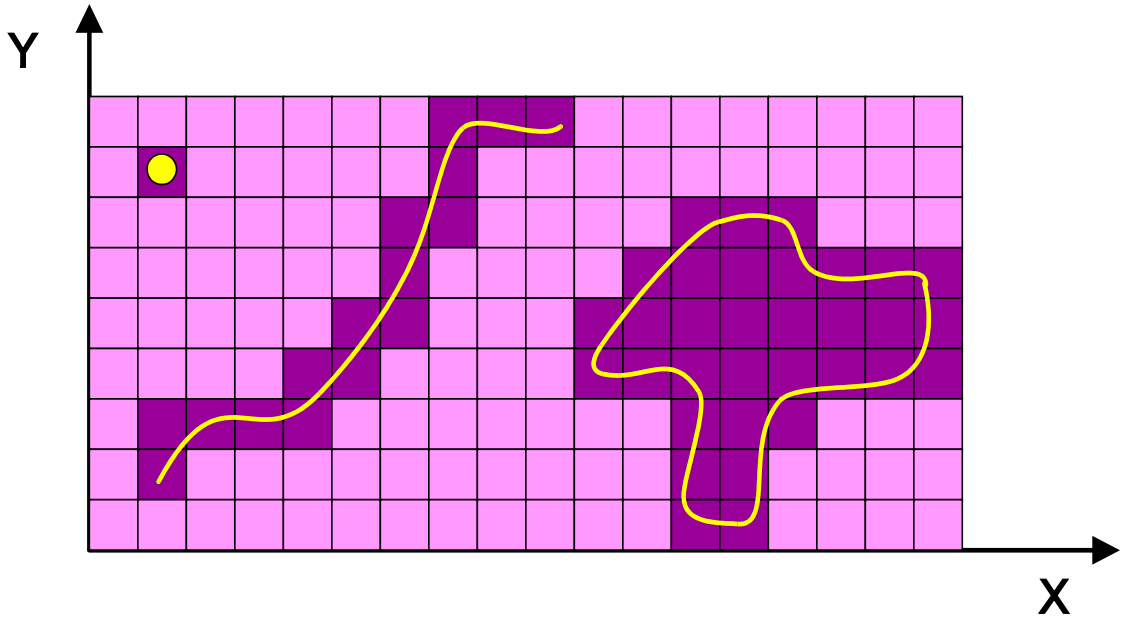


- Níveis de abstração
  - Ontologias (conceitos do mundo real): lote, tipo de solos
  - Formal: distribuições (campos) X entidades (objetos)
  - **Estruturas de dados: matrizes, vetores**
  - Implementação: código em linguagem de computador

Fonte: DPI / INPE

Universo Estrutural:
Métodos de Representação
(Modelo Lógico)
de Dados Espaciais

Método de Representação Matricial

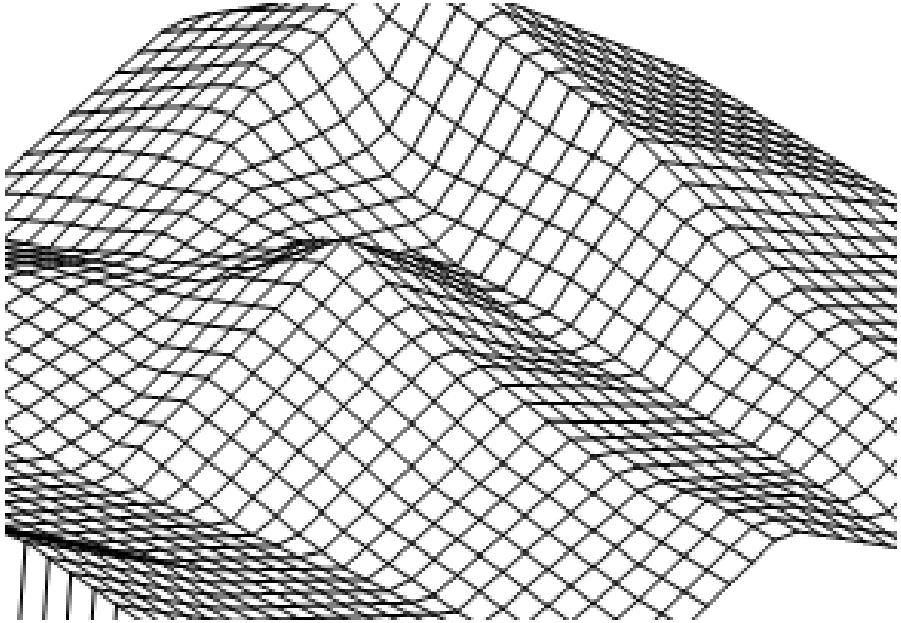


0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0
0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	1
0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	1	1
0	1	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0

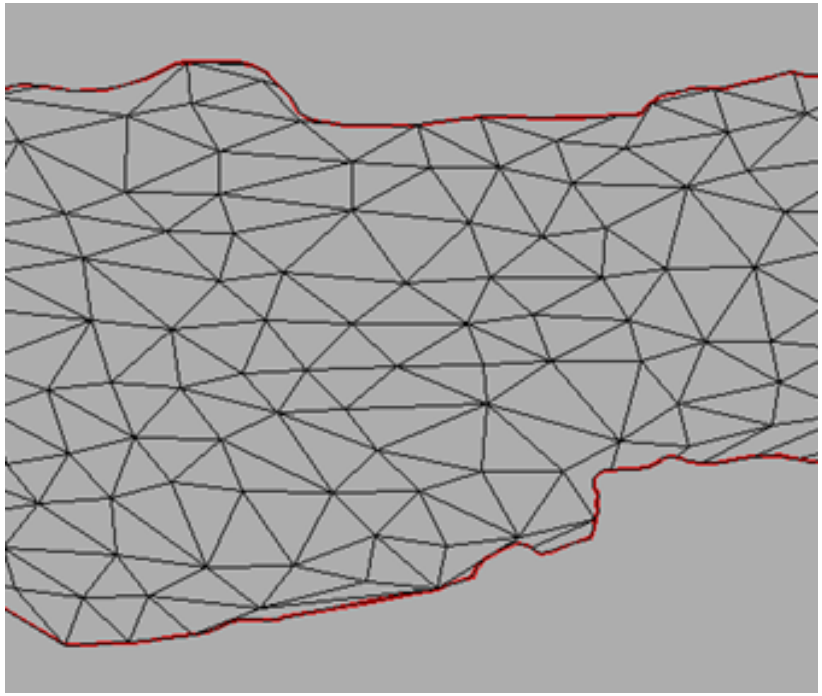
Representação do MDT

Representação gráfica

Rede Quadrada



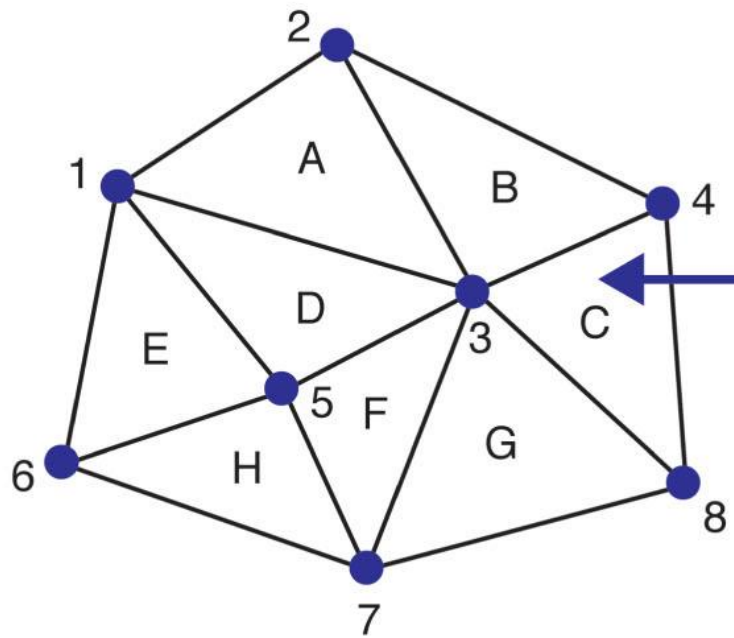
Rede triagular



© Copyright LTG/PTR/EPUSP

The topology of a TIN

A TIN is a topologic data structure that manages information about the nodes that comprise each triangle and the neighbors to each triangle



Triangle	Node list	Neighbors
A	1, 2, 3	-, B, D
B	2, 4, 3	-, C, A
C	4, 8, 3	-, G, B
D	1, 3, 5	A, F, E
E	1, 5, 6	D, H, -
F	3, 7, 5	G, H, D
G	3, 8, 7	C, -, F
H	5, 7, 6	F, -, E

Triangles always have three nodes and usually have three neighboring triangles. Triangles on the periphery of the TIN can have one or two neighbors.

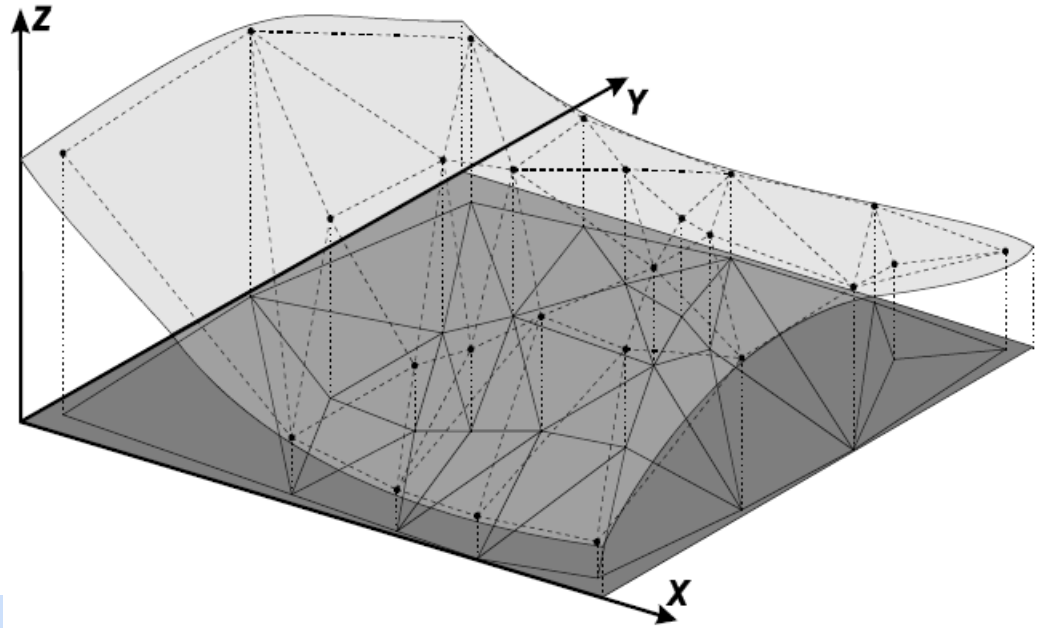
(Source: after Zeiler 1999)

Tipos de Rede

Rede irregular triangular
(*"TIN" – Triangular Irregular Network*)

Superficie onde cada polígono, que forma uma face do poliedro, é um triângulo.

- ✓ Melhor definição do relevo;
- ✓ Menor quantidade de amostragem;
- ✓ Utiliza os próprios dados de amostragem;



© Copyright LTG/PTR/EPUSP

Modelo (visão) de Campos contínuos: Método de Representação Matricial (Raster)

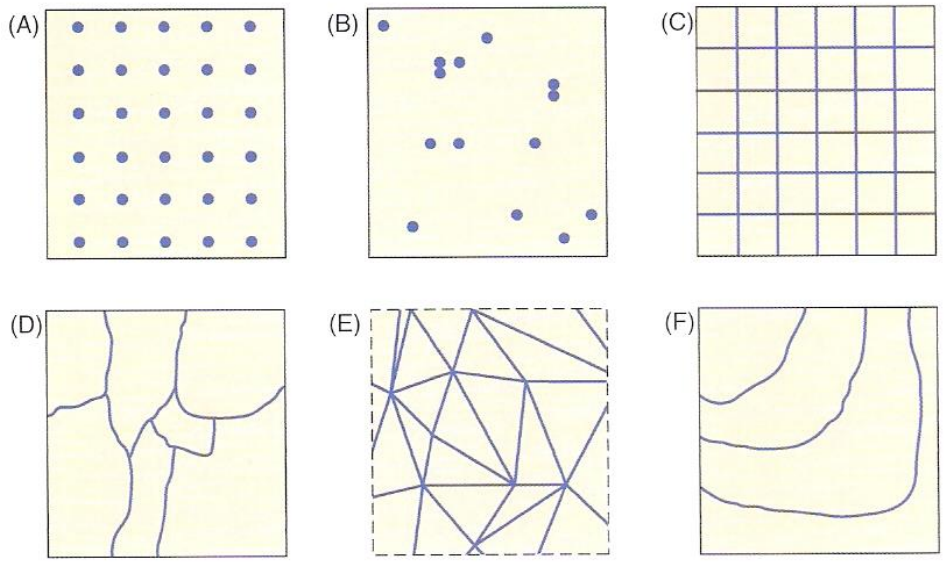
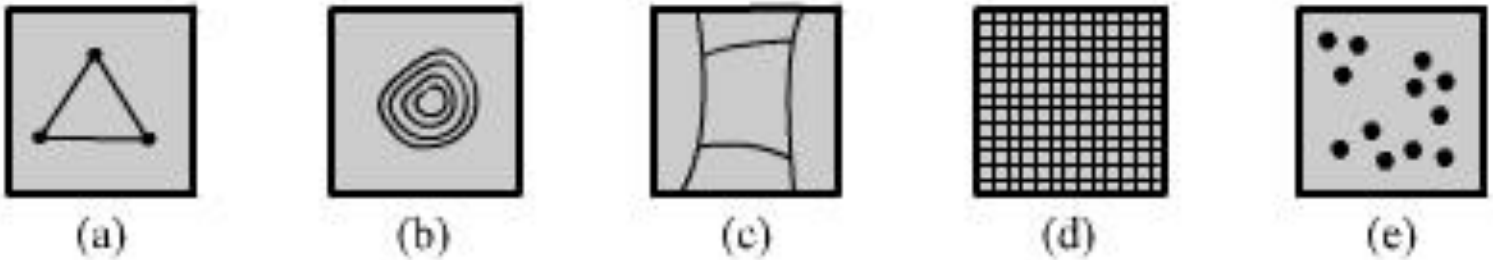


Figura 3.12 As seis representações aproximadas de um campo contínuo utilizado em SIG. (A) Pontos amostrais regularmente distribuídos. (B) Pontos amostrais irregularmente distribuídos. (C) Células retangulares. (D) Polígonos de formas irregulares. (E) Rede triangular irregular, com variação linear sobre cada triângulo (o modelo de Rede Triangula Irregular ou TIN; a moldura é mostrada tracejada neste caso porque porções não mostradas de triângulos completos a ultrapassam) (F) Polilinhas representando curvas de nível (ver a discussão dos mapas isopléticos no Quadro 4.3). (Cortes U.S. Geological Survey)

Geo-Campo

Refere-se a qualquer posição no espaço geográfico considerado, corresponde a algum valor da variável representada.

O modelo Geo-OMT possui **cinco subclasses do tipo Geo-Campo**:

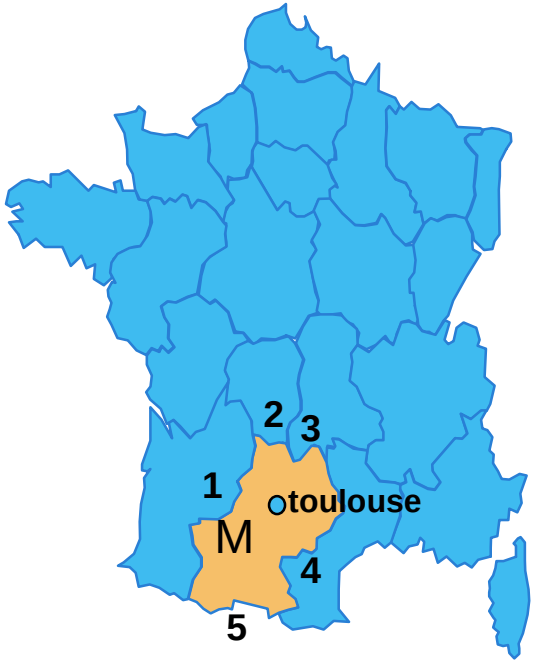


(a) Rede Triangular Irregular - **Exemplo:** TIN
(b) Isolinhas - **Exemplo:** Nível
(c) Polígonos Adjacentes - **Exemplo:** Divisão de Bairros
(d) Tesselação - **Exemplo:** Imagens de Satélite
(e) Amostragem - **Exemplo:** Pontos cotados

Método de Representação Vetorial

- Componentes
  - ponto, linha, região
  - região M={1,2,3,4,5}

- Topologia
  - relação espacial entre objetos
  - Toulouse fica na região M

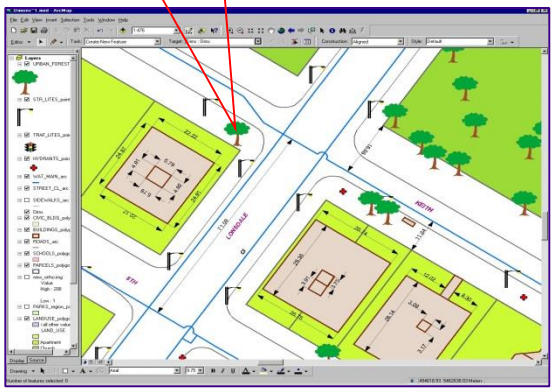


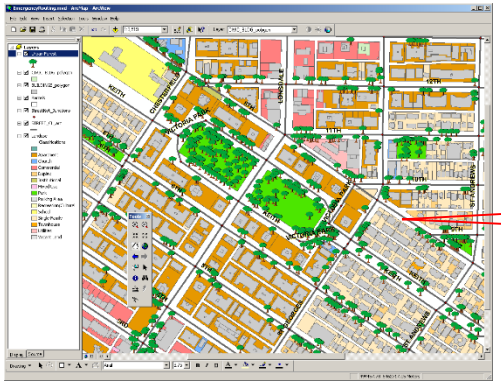
Fonte: adaptado de DPI / INPE

Método de Representação Vetorial – exemplos

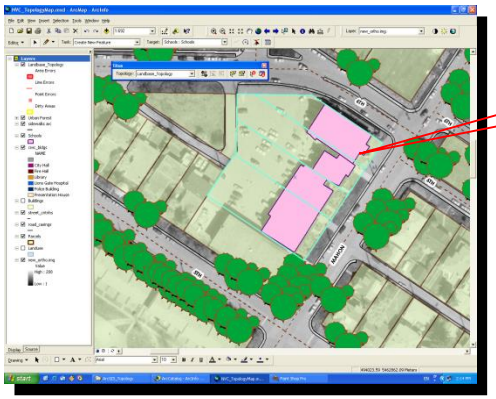
Os entes vetoriais (ponto, linha e polígono) simbolizam o mundo real.

Ponto = Árvore





Linhas = Ruas



Polígonos = Lotes

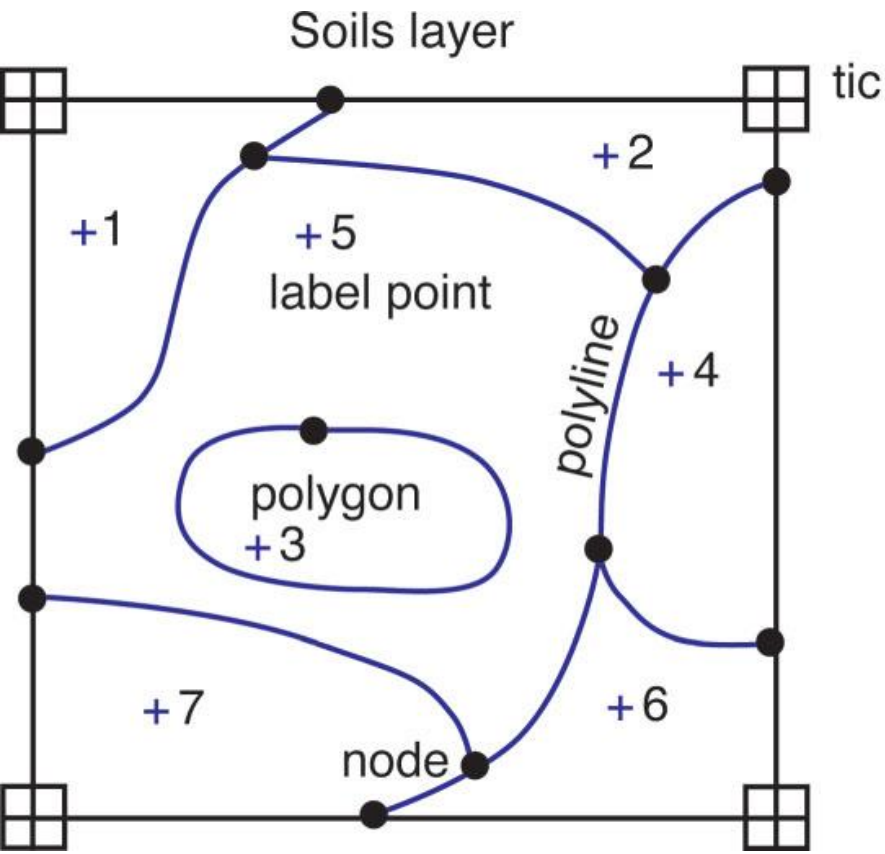
Different representational models of the same area

in Colorado, USA

(A)aerial photograph

(B) vector objects, some digitized from the photograph

An example of a georelational polygon dataset



Soils attributes

ID	Soil	Class	Suitability
1	A3	113	HIGH
2	C6	95	LOW
3	B7	212	MODERATE
4	B13	201	MODERATE
5	Z22	86	LOW
6	A6	77	HIGH
7	A1	117	LOW

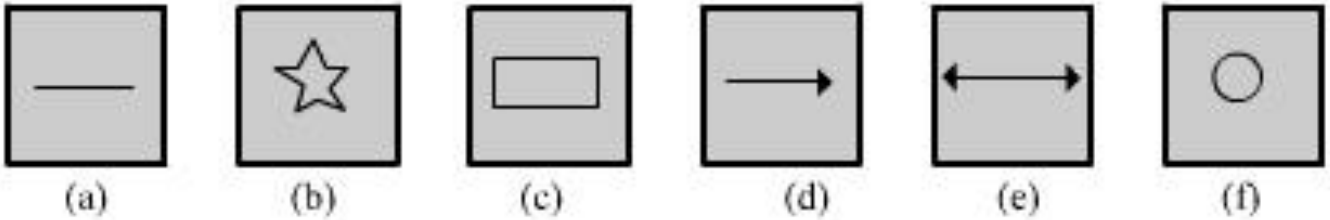
Each of the polygons is linked to a row in an RDBMS table. The table has multiple attributes, one in each column.

(Source: after ESRI 1997)

Geo-Objeto

O modelo Geo-OMT possui duas **subclasses do tipo Geo-Objeto**:

- Geo-Objeto com Geometria;
- Geo-Objeto com Geometria e Topologia.



GEO-OBJETO com Geometria

(a) Linha - **Exemplo:** Muro

(b) Ponto - **Exemplo:** Árvore

(c) Polígono - **Exemplo:** Lote

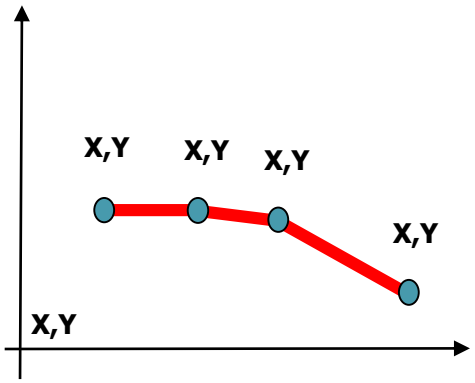
GEO-OBJETO com Geometria e Topologia

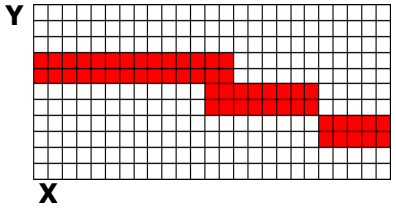
(d) Linha Uni-direcionada - **Exemplo:** Trecho de Rede de Esgoto

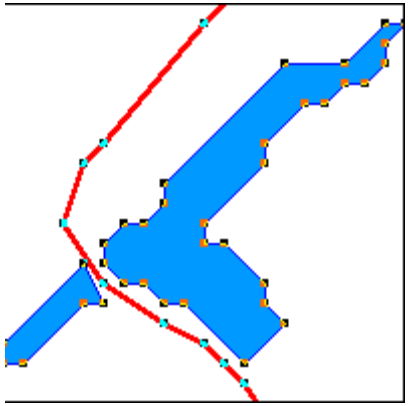
(e) Linha Bi-direcionada - **Exemplo:** Trecho de Rede de Água

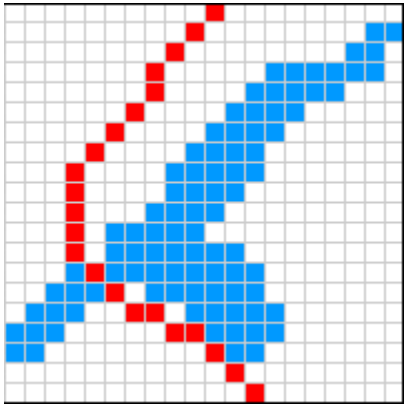
(f) Nó - **Exemplo:** Entroncamento de Rodovias

Vetorial x Matricial









Questões para estudo

LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2013, 539p

Capítulo 3: A Representação Geográfica [2 - PRINCÍPIOS]

Pág 97

Básicas: Questões 1 e 2

Maior tempo para pesquisa (opcionais): Questões 3 e 4

Questões para estudo

Questão 2 (Complemento) Use a Tabela 3.3 para avaliar a adequação das duas formas de representação para algumas aplicações de SIG.

Justifique as afirmações abaixo:

a) para a produção de cartas e em operações que requeiram maior precisão, a representação vetorial é mais adequada.

b) as operações de álgebra de mapas são mais facilmente realizadas no formato matricial.

Tabela 3.3 Vantagens relativas das representações matricial e vetorial

Aspecto	Matricial	Vetorial
Volume de dados	Depende do tamanho da célula	Depende da densidade de vértices
Fontes de dados	Sensoriamento remoto, imagens	Dados sociais e ambientais
Aplicações	Recursos naturais, ambientais	Sociais, econômicas, administrativas
Software	SIG matricial, processamento de imagens	SIG vetorial, cartografia automatizada
Resolução	Fixa	Variável

Técnicas de Modelagem de Dados Espaciais

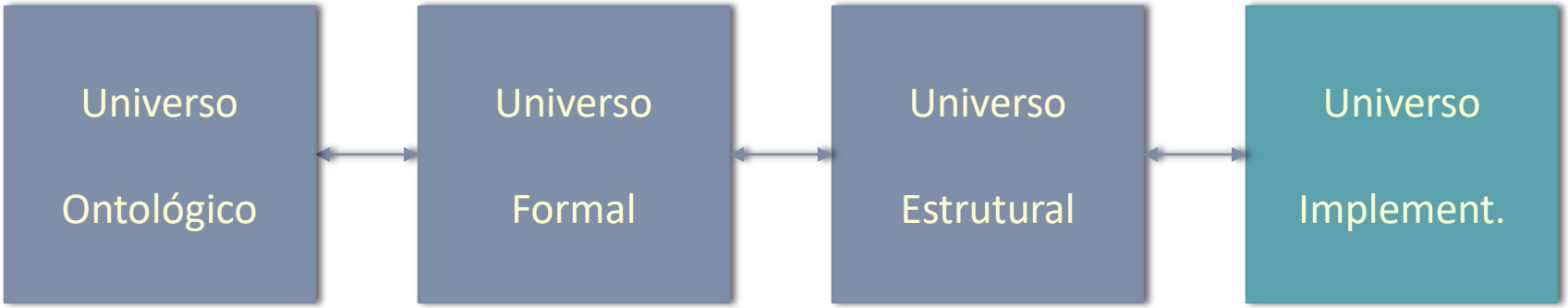
Agenda 1

Bibliografia

- Modelagem de dados geográficos:

- LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2011, 539p
  - **Capítulo 8: Modelagem de Dados Geográficos [TÉCNICAS]**
- CASANOVA, Marco; CÂMARA, Gilberto; DAVIS, Clodoveu; VINHAS, Lúbia; QUEIROZ, Gilberto Ribeiro. **Bancos de Dados Geograficos**. 1ª. ed. Curitiba: Editora Mundo Geo, 2005. 504 p. (Disponível em: http://www.dpi.inpe.br/gilberto/livro/bdados/index.html)
  - http://www.dpi.inpe.br/livros/bdados/cap3.pdf

Processo de Representação Computacional



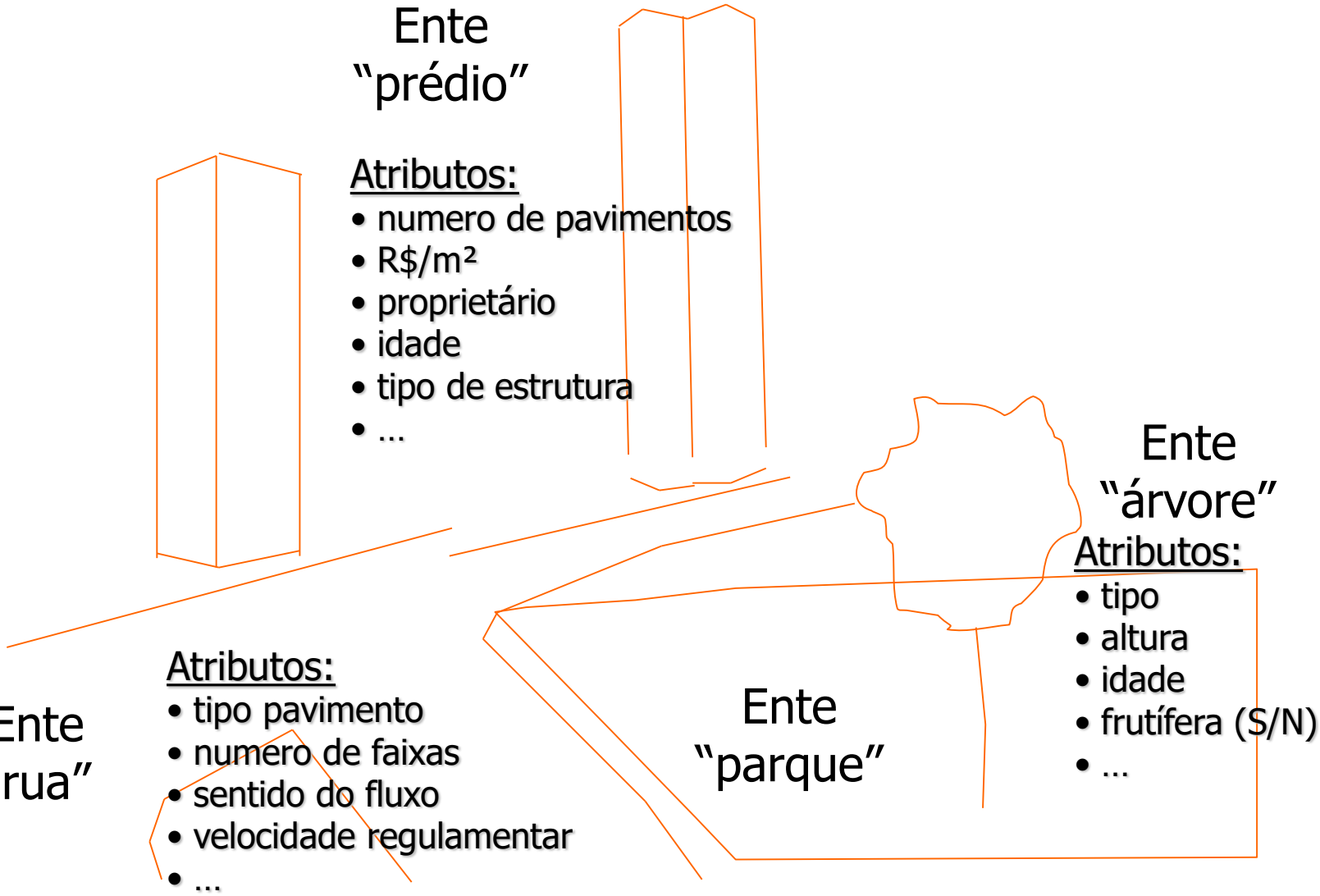
- Níveis de abstração
  - Ontologias (conceitos do mundo real): lote, tipo de solos
  - Formal: entidades (objetos) x distribuições (campos)
  - Estruturas de dados: matrizes, vetores
  - **Implementação: em Banco de Dados**

Fonte: DPI / INPE

Representação Computacional

Fonte: Carlos Vieira (Geocapacitar/UFV)

Representação Computacional: entes e atributos



Fonte: Carlos Vieira (Geocapacitar/UFV)

Os processos (estágios) da Modelagem de Dados Espaciais

1. Modelagem de Dados Conceitual
   - Universo Formal:
     - ✓ **Modelos Conceituais** Fundamentais (Visões)
     - ✓ Objetos Discretos e Campos Contínuos

2. Modelagem de Dados Lógica
   - Universo Estrutural:
     - ✓ **Métodos de Representação [Modelos Lógicos]**
     - ✓ **Matricial e Vetorial**

3. Modelagem de Dados Física
   - Universo de Implementação:
     - ✓ **Técnicas de Modelagem**
     - ✓ MOO, OMT (UML) e Geo-OMT (OMT-G)

Modelos de Dados

Os modelos de dados anteriores são centrados na geometria

- ✓ modelam o mundo como pontos, linhas áreas, TIN, ou raster/células matriciais.

Essa abordagem pode apresentar diversas limitações para a modelagem de dados geográficos.

Fonte: Longley et al., 2013

Modelos de Dados

Os sistemas geográficos normalmente contém muitos tipos diferentes de entidades com um grande número de propriedades, relações complexas e comportamento sofisticado.

Fonte: Longley et al., 2013

Modelos de Dados

A modelagem de tais entidades como simples tipos geométricos é excessivamente simplista e não suporta com facilidade as características sofisticadas que são necessárias em algumas análises.

Fonte: Longley et al., 2013

Modelo de Dados de Objeto

Para tentar resolver esses problemas foram criados os Modelos de Dados de Objetos

- ✓ permitem a modelagem da riqueza total dos sistemas geográficos
- ✓ de uma forma integrada em um SIG

O foco de um Modelo de Dados de Objetos para um SIG é a coleção de objetos geográficos e seus relacionamentos

Fonte: Longley et al., 2013

Modelo de Dados de Objeto

(Cap 10):

O resultado de um exercício de modelagem de dados é um Projeto de Banco de Dados Físico.

Este projeto irá incluir a especificação de todos os tipos de dados e relacionamentos, bem como a configuração real do banco de dados necessário para armazená-los.

Fonte: Longley et al., 2013

Modelagem de Dados

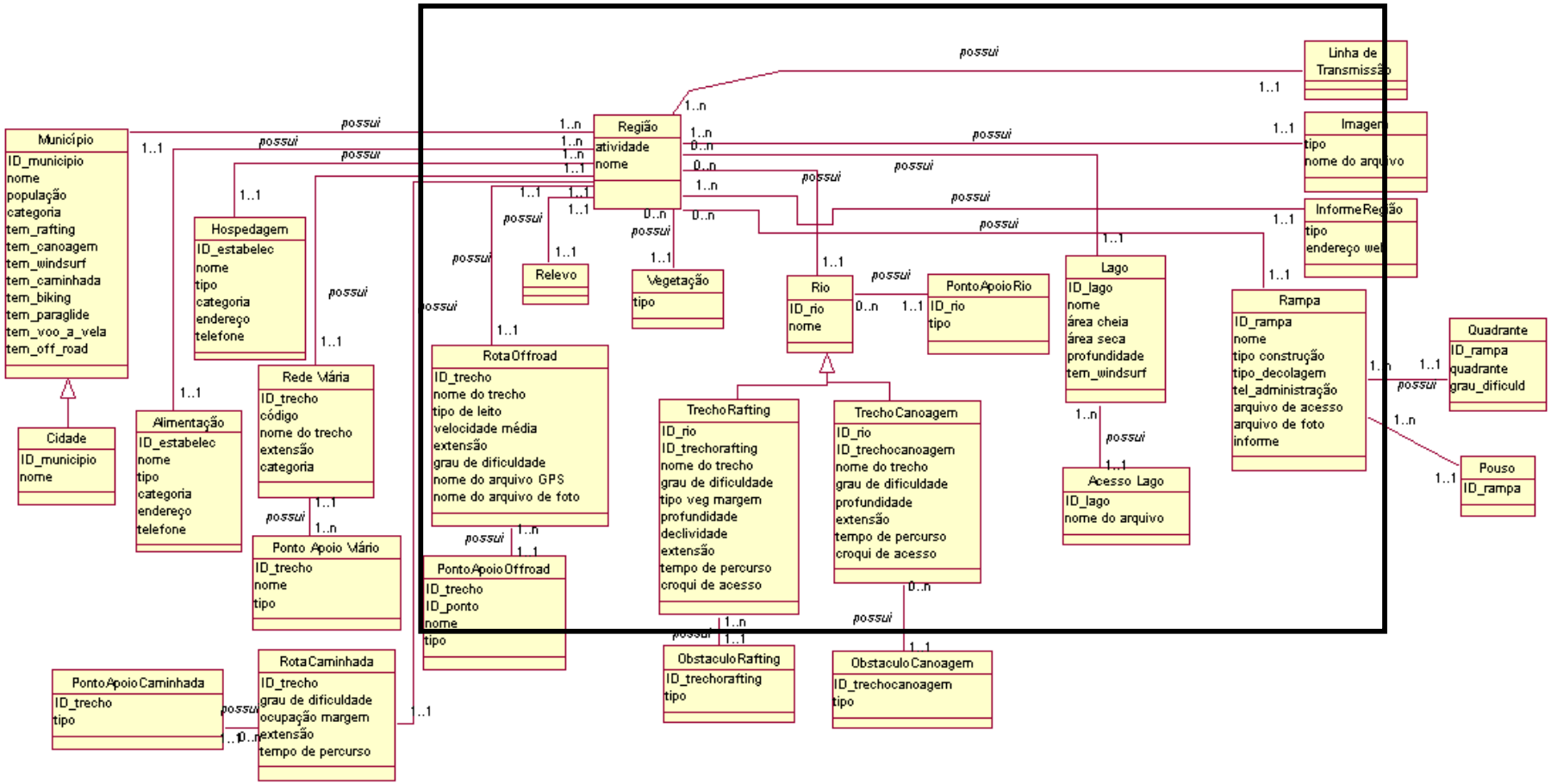
- Está relacionado com a maneira pela qual o observador vê o mundo.
- As informações são representadas graficamente.
- Sem detalhes de implementação
  - ✓ ou descrição de procedimentos.
- Existem diferentes metodologias usadas na modelagem de dados conceitual (lógica):
  - ✓ MER (Modelo Entidade-Relacionamento)
  - ✓ OMT (Object Modeling Technique)
    - UML (Unified Modeling Language)

Técnicas de Modelagem de Dados Espaciais

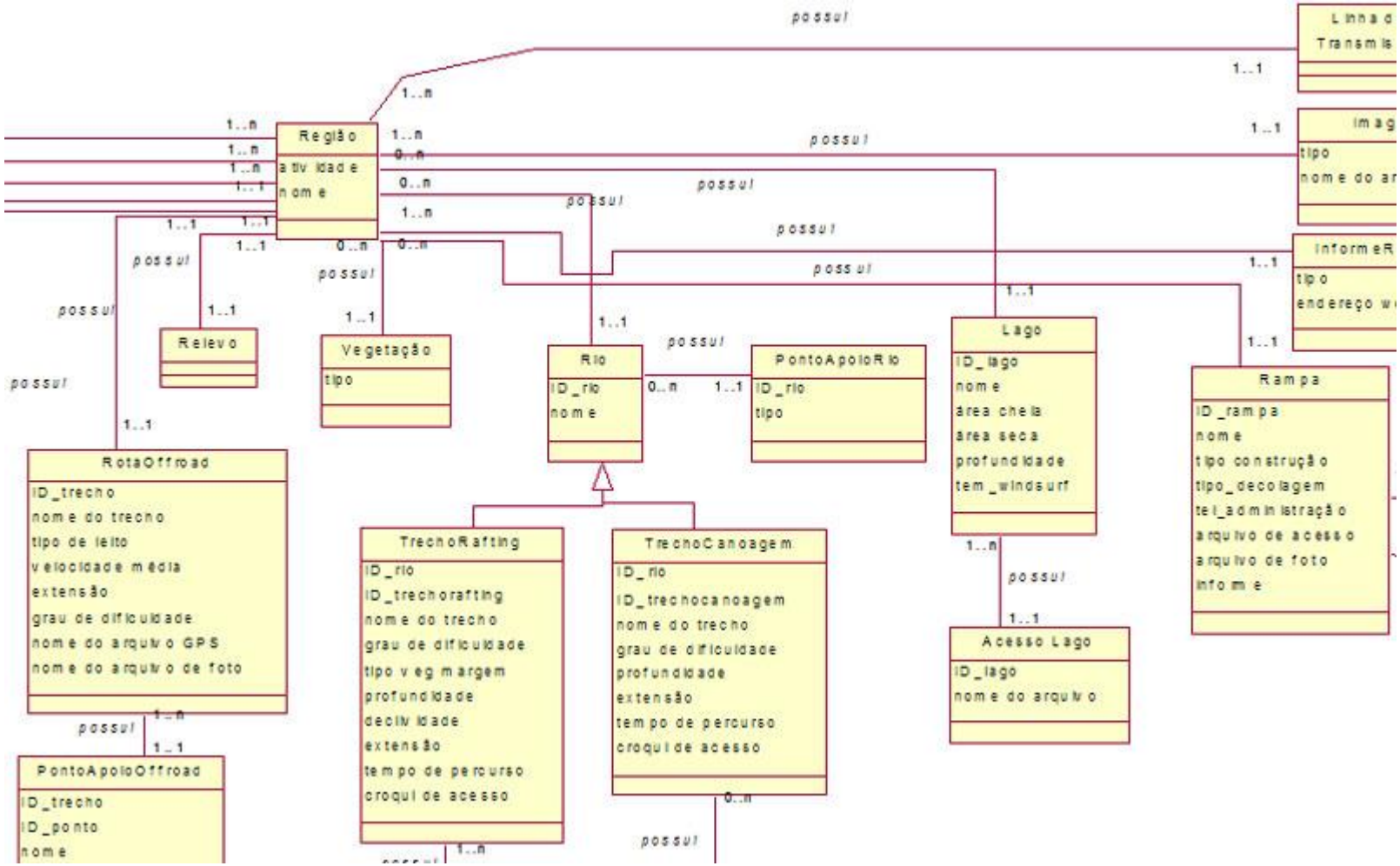
Parte I – Modelo Orientado a Objetos (MOO)

UML - *Unified Modeling Language*

Modelagem da Serra da Mantiqueira



Modelagem da Serra da Mantiqueira



Modelo Orientado a Objetos – MOO

Modelo Orientado a Objeto surgiu a partir da programação Orientada a Objetos – OO, considerando, entre outros, os conceitos de: herança, encapsulamento e abstração.

As linguagens de programação OO surgiram da necessidade de desenvolvimento de arquiteturas de sistemas que possuem dados espalhados em uma rede dispersa geograficamente e conectada por meio de canais de comunicação.

MOO – **Objeto ("Ente")**

O Objeto é uma **unidade atômica básica em um Modelo de Dados de Objeto e compreende todas as propriedades que definem o estado de um objeto, juntamente com os métodos que definem seu comportamento**.

O tipo do Objeto possui uma definição interna e externa (Encapsulamento):

- **Definição Interna**: refere-se à operação do Objeto, o que não é visível ao usuário.
- **Definição Externa**: refere-se ao **conjunto de operações, propriedades, atributos e exceções** que os usuários podem ver e acessar.

Uma operação é um comportamento abstrato do Objeto, definido por uma linguagem de programação (específica);

O Objeto pode possuir herança: os tipos possuem propriedades e operações que poderão ser transmitidas aos subtipos de uma classe.

MOO – Objetos e Classes (1)

São **agrupamentos de Objetos** em Classes de Objetos ou apenas Classe.

- ✓ É um template (uma descrição) para criar objetos.

Descrevem a estrutura (estática) de um sistema.

Um objeto é uma instância de uma Classe.

Um grupo (uma coleção) de Objetos com:

- Propriedades semelhantes (atributos);
- Mesmo comportamento (operações);
- Mesmo relacionamento com outros Objetos;
- Mesma semântica (significado).

MOO –Objetos e Classes (2)

Todo Objeto possui

- uma identidade
- um estado
- um comportamento

Classes

- Conjunto de objetos que compartilham estrutura, relações e comportamento
- **Representam algo do mundo real**

MOO – Classes

Onde estão as classes?

- Nos sujeitos da descrição do problema
  - Ex.: O Carro é alugado
- Nas situações (tipos)
  - Ex.: O aluguel do carro
- Nos papéis (especialização)
  - Ex.: O funcionário é um gerente
  - Ex.: O carro é bicombustível (flex)

MOO – **Diagrama(s)** de Objetos (1): Tipos

É uma **notação gráfica formal** para a modelagem de objetos e seus relacionamentos.

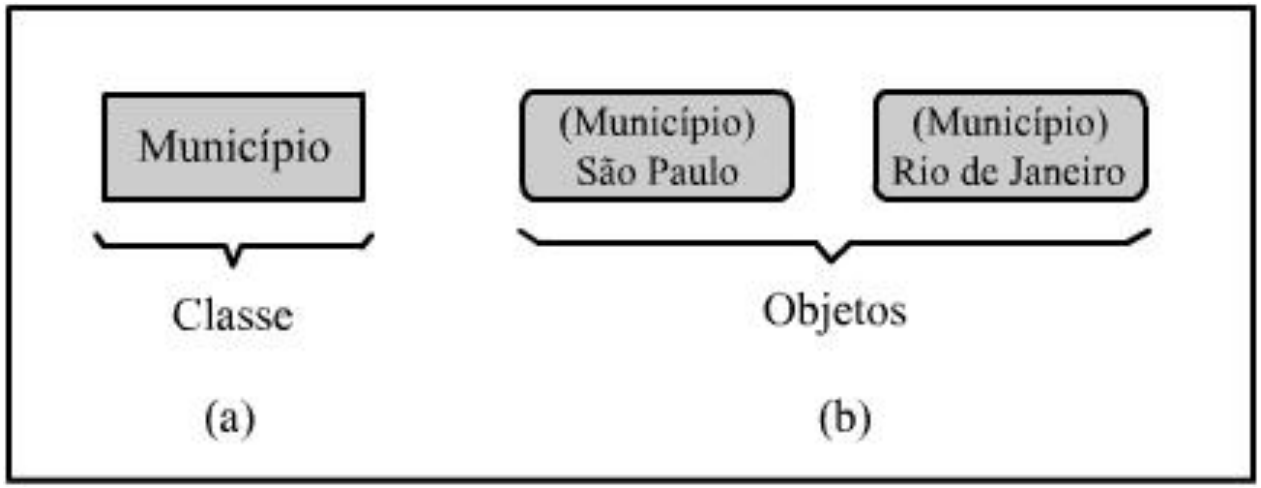
Diagramas de objetos estruturam a notação gráfica formal

para a modelagem de objetos e seus relacionamentos.

Existem alguns tipos de diagramas de objetos:

- Diagramas de Classes (**Estáticos**): é um esquema ou um modelo que descreve instâncias possíveis de dados.
- Diagramas de Instâncias (Dinâmicos): descrevem como os objetos de um determinado conjunto se relacionam entre si
  - Exemplo: Seqüência, Colaboração.

MOO – Diagrama de Objetos (2): Classes



A diferença das representações está baseada no modelo OMT, onde diagramas de classes são retângulos com seu nome em negrito, e os diagramas de instâncias são retângulos com cantos arredondados.

MOO – **Atributos** (1)

São dados associados a uma Classe.

São as **propriedades** dos objetos definidas por uma variedade de valores.

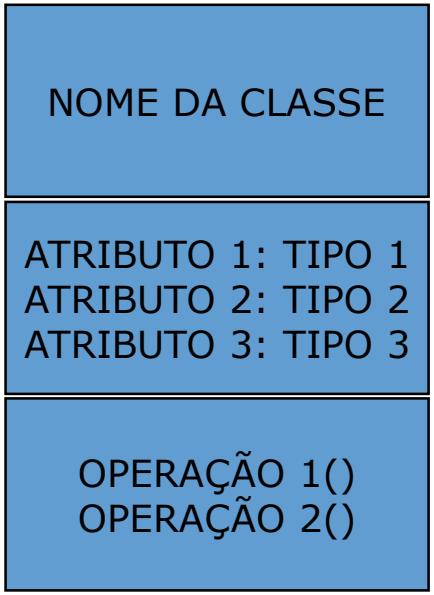
Cada Atributo possui um valor para cada instância de um objeto.

São alterados pelas Operações (preferencialmente internas às Classes).

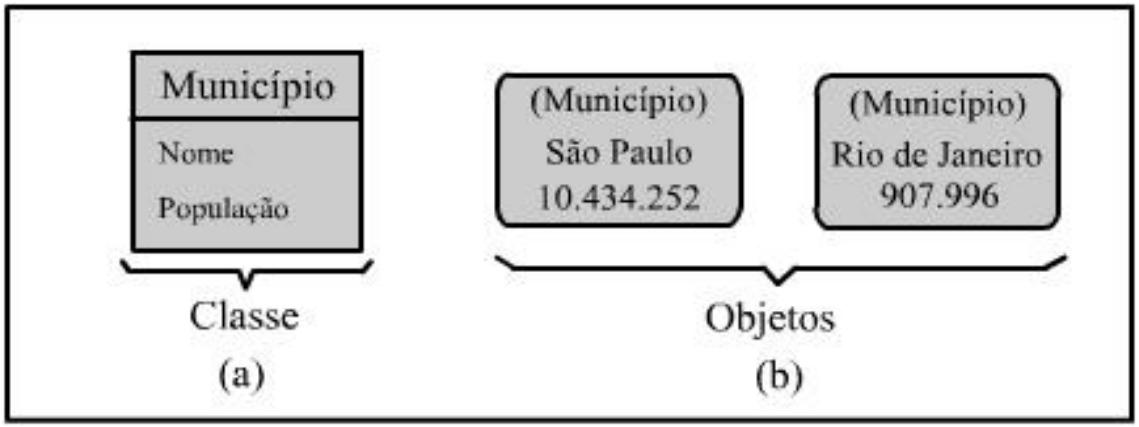
No Diagrama os Atributos são mostrados no **segundo** bloco do quadro de uma classe.

MOO – Atributos (2)

MODELO



EXEMPLO



Cada objeto possui um identificador único e não nulo que o distingui dos demais objetos.

Esse identificador é gerado automaticamente nas modelagens orientadas a objeto.

MOO – Atributos (3)

São bons Atributos

- Nome do funcionário
- Cor do carro
- Saldo da conta corrente

Identificando os atributos

- São os adjetivos
- Aquilo que caracteriza uma Classe

Evitar atributos desnecessários

MOO – Atributos (4):
Banco de Dados - Modelo Relacional

- **Atributo**
  - É o nome dado a cada **coluna** de uma tabela.
  - Pode estar sujeito a um **domínio de valores.**

Atributos

Cidade

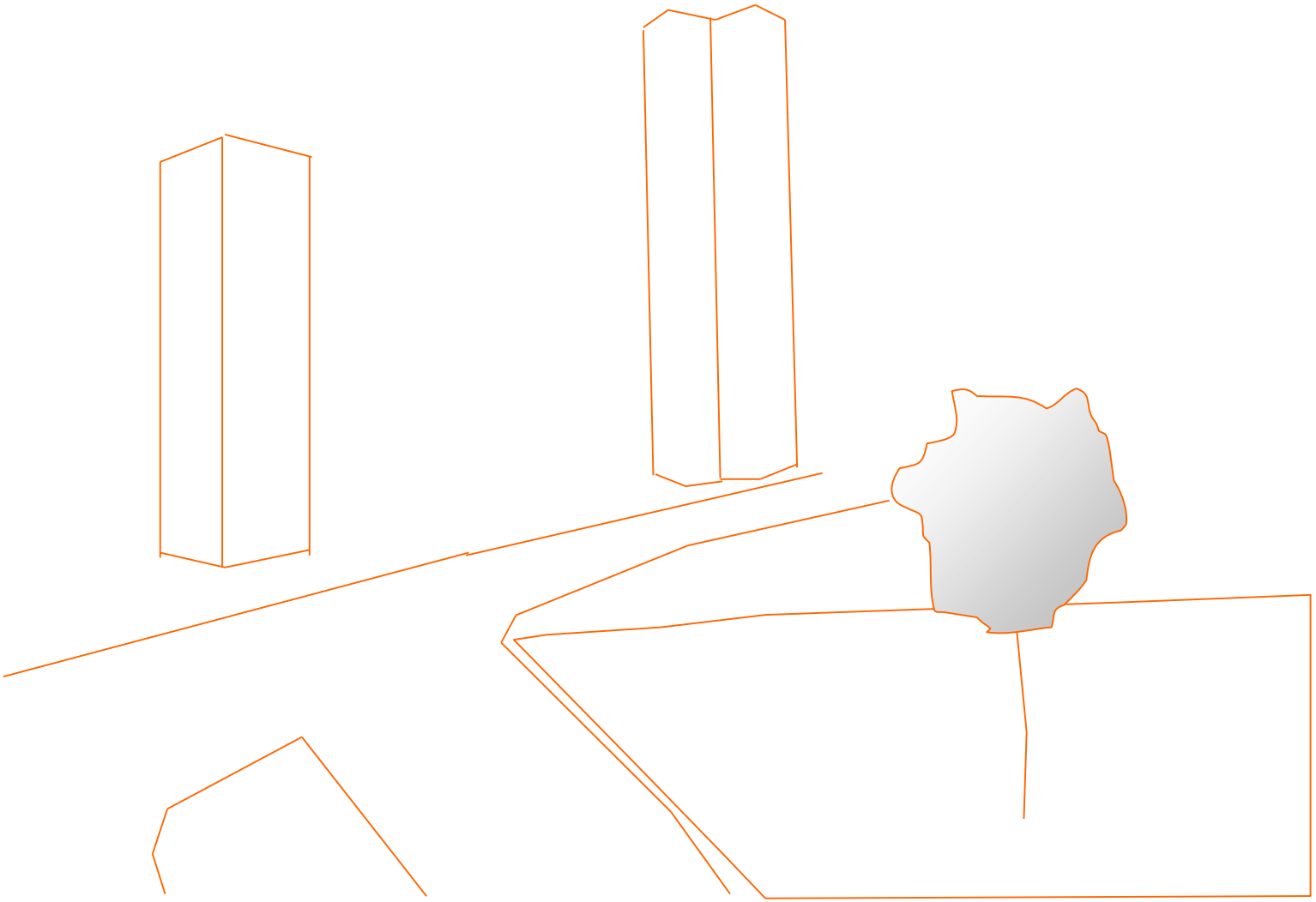
CID_CODIGO	CID_NOME	CID_POPULACAO
1	SAO PAULO	1000000
2	ARACARIGUAMA	15000
3	SAO ROQUE	70000
4	ITAPEVI	900000
5	BARUERI	120000

Tuplas

Representação Computacional

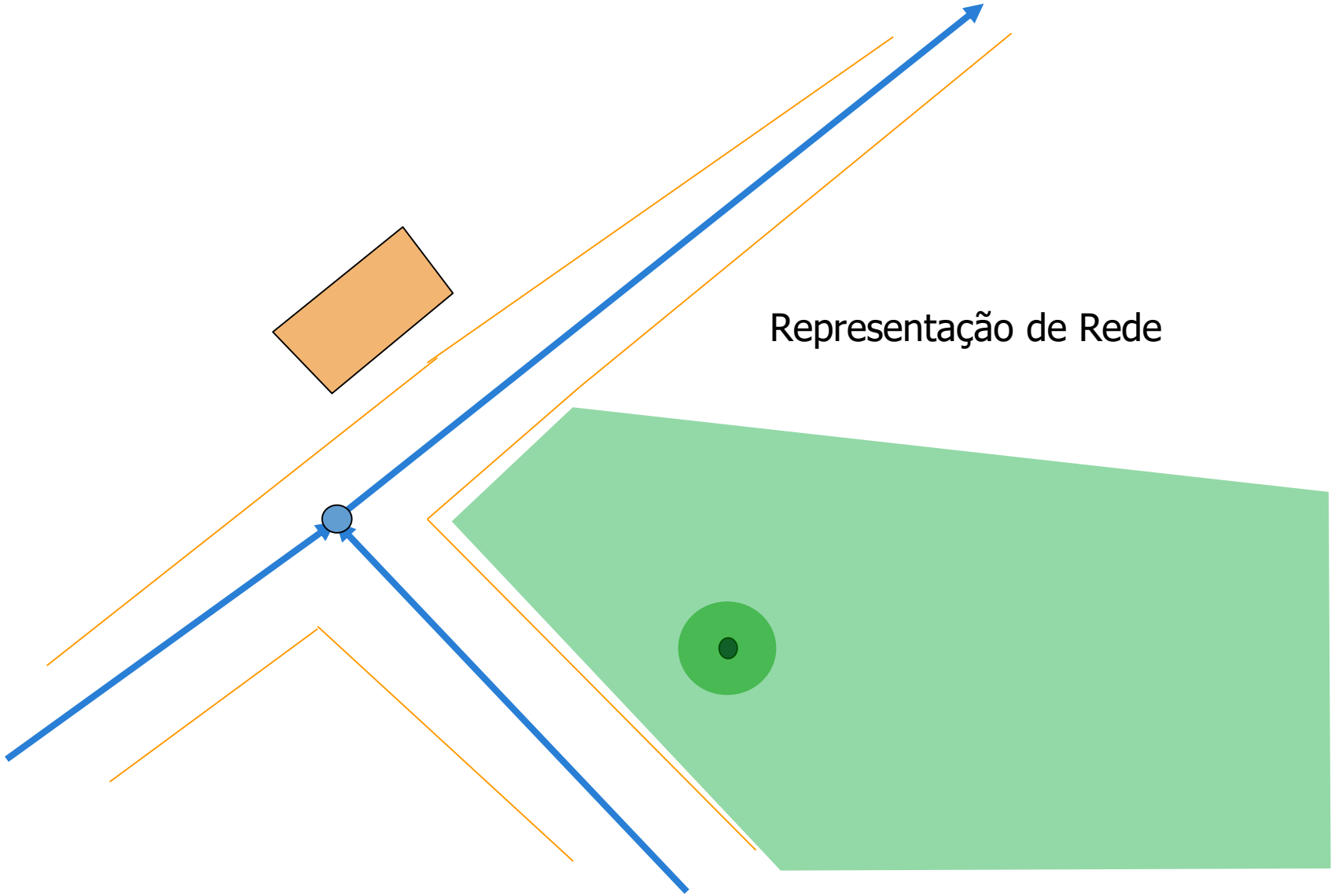
Fonte: Carlos Vieira (Geocapacitar/UFV)

Representação Computacional: 3D



Fonte: Carlos Vieira (Geocapacitar/UFV)

Representação Computacional: 2D



Fonte: Carlos Vieira (Geocapacitar/UFV)

MOO – **Operações e Métodos** (1)

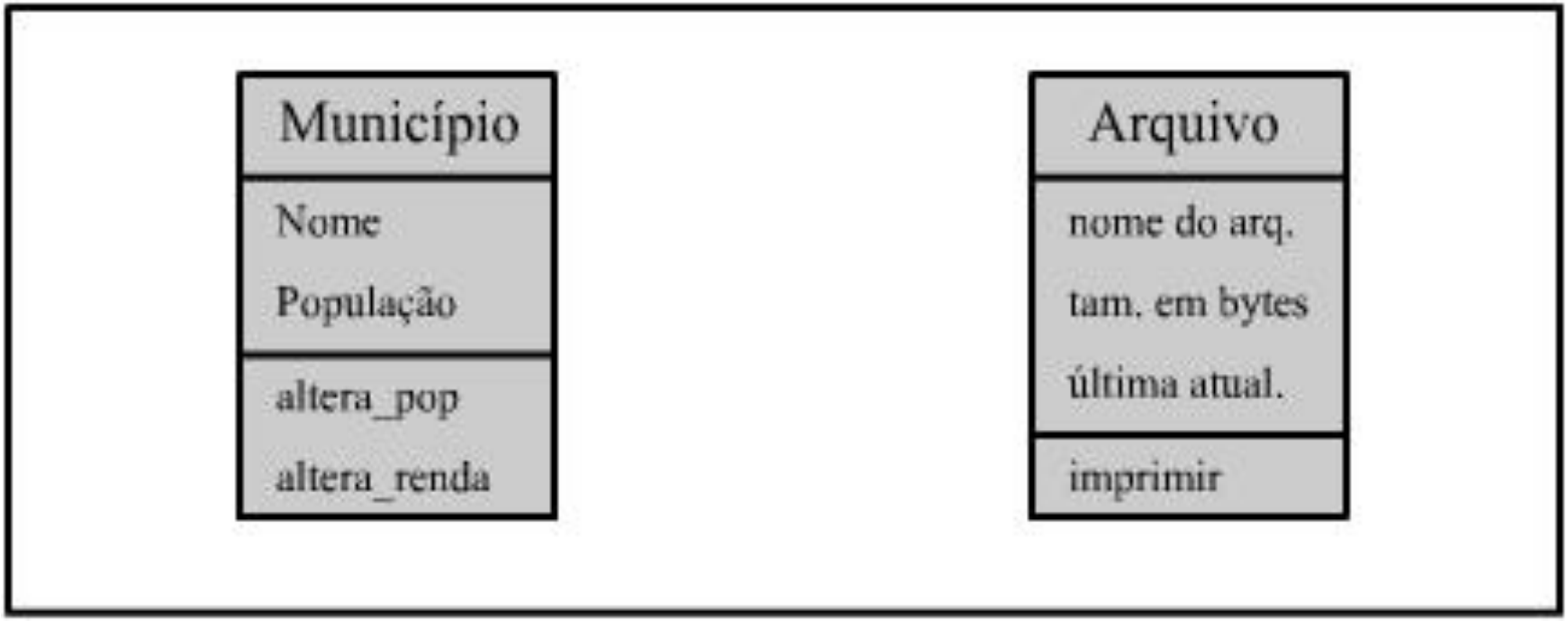
Uma **Operação** é um serviço que se pode requisitar para uma instância de uma classe ou objeto.

- Uma função que pode ser aplicada ao objeto ou que o objeto deva cumprir.

É implementada a partir de um **Método**, que é o código executável da função requisitada, localizada em uma **Interface**.

Coleções de Operações podem ser organizadas em **Interfaces**, permitindo estas sejam utilizadas por diferentes Classes de Objetos.

MOO – **Operações e Métodos** (2)



As operações e os métodos são apresentados no **terceiro** bloco do quadro de uma Classe.

MOO – **Operações** (3)

Identificando as Operações

- São ações na descrição do sistema
- Funções que o sistema deve realizar e quem as realiza

Exemplos

- <u>Alugar</u> carro
- <u>Retornar</u> livro
- <u>Solicitar</u> confirmação

MOO – Relacionamentos

Os Relacionamentos entre Classes no MOO são estabelecidos por meio de Ligações e Associações.

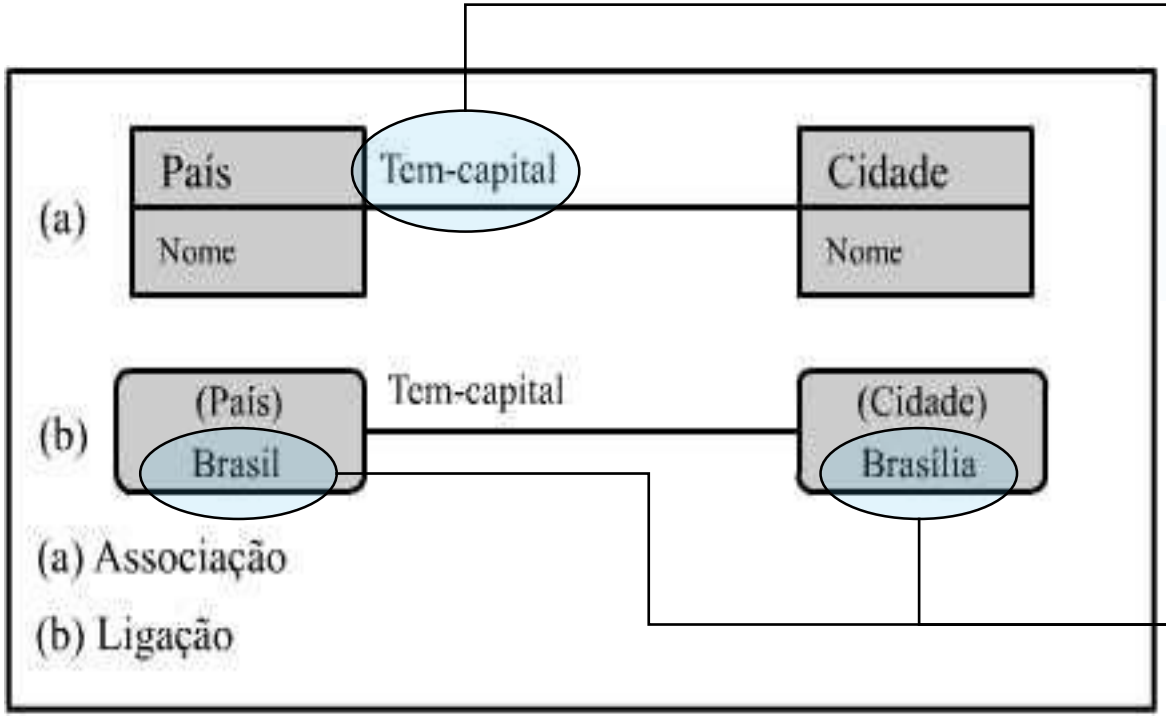
A **Ligação** é a conexão física ou conceitual entre instâncias de objetos.

A **Associação** é o grupo de ligações com estruturas semânticas comuns.

Uma associação descreve um conjunto de potenciais ligações da mesma maneira que uma classe descreve um conjunto de potenciais objetos.

Três tipos de relacionamentos são frequentemente utilizados em Modelos de Dados de Objetos Geográficos: Topológicos, Geográficos (operadores geográficos) e gerais

MOO – Ligações e Associações



A associação mostra como um objeto de uma classe é relacionado a outro objeto. Relacionamento semântico:

A ligação é uma instância de uma associação

A notação gráfica para associações e ligações é uma linha que liga as classes.

MOO – Atributo ou Associação ???

- É uma dúvida muito comum
- Semanticamente não existe diferença
- Usar atributo quando:
  - A Classe não tem significado isolado
  - Não há operações isoladas
  - OPTAR PELA SIMPLICIDADE

MOO – **Multiplicidade**

A Multiplicidade é uma propriedade do relacionamento ou da associação.

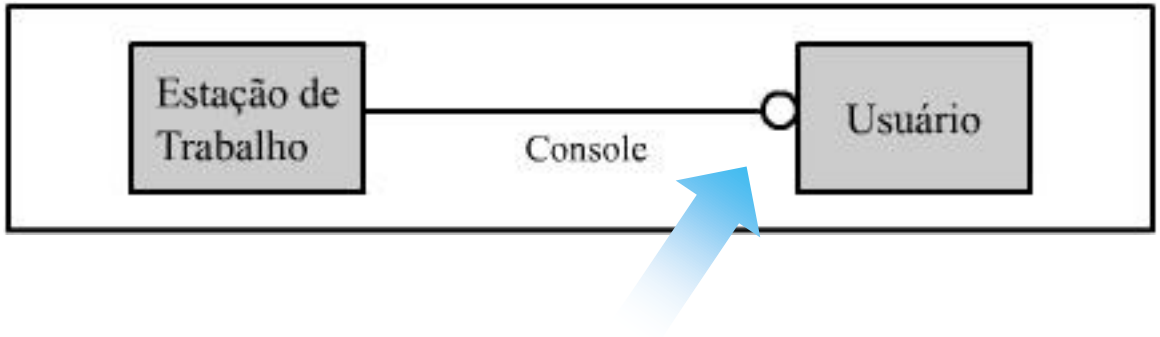
A representação da multiplicidade em MOO é chamada **terminadores**.

Indica a **Cardinalidade (ocorrências)** que a classe pode assumir.

As multiplicidades podem ser entendidas como “um” ou “muitos”.

Exemplos: 0 .. 1 0 .. *

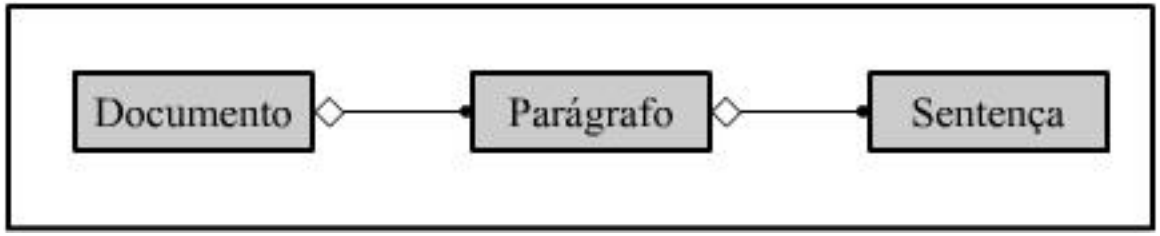
Uma linha sem símbolos indica uma associação um-para-um.



MOO – **Agregação** (1)

Agregação é um **relacionamento de complementaridade**, ou seja, um objeto pode agregar outros objetos ou classes de objetos.

É um Relacionamento **da parte com o todo** entre duas classes.



AGREGAÇÃO COMPARTILHADA: O Documento agrega Parágrafos e Sentenças. Por sua vez, o Parágrafo agrega as Sentenças.

AGREGAÇÃO COMPOSTA: Os parágrafos são propriedades do documento – PRINCÍPIO DA PROPRIEDADE

MOO – Agregação (2)

Quando usar Agregação ?

- Quando se encontra uma composição ou decomposição na descrição do sistema
- Quando há frases do tipo:
  - A é formado de B
  - B é parte de A
  - A possui B
- Há uma dependência de existência
- B não existe sem A

MOO – **Herança** (Generalização e Especialização) [1]

É o **relacionamento hierárquico** entre um elemento mais geral e um elemento mais específico, ou seja, define hierarquia de herança entre Classes:

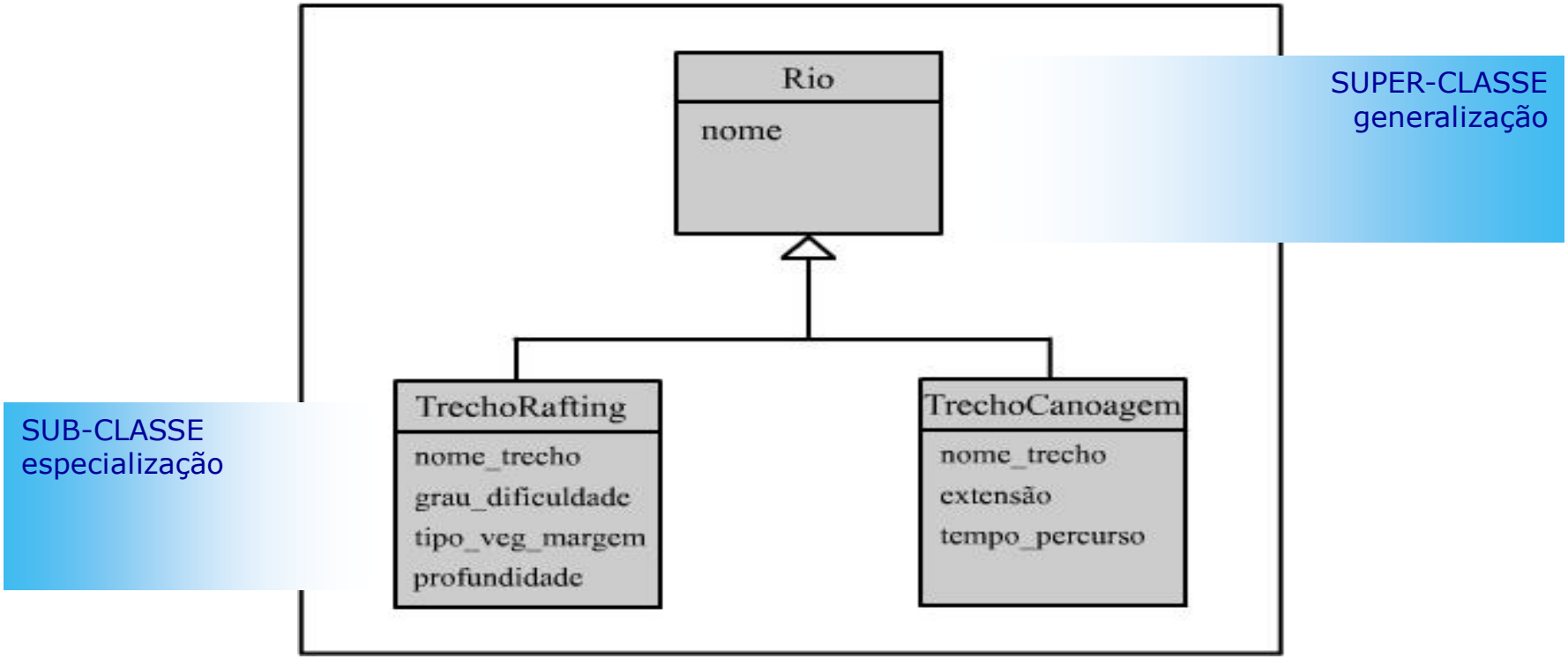
- Classes herdam elementos de outras Classes.

Uma Classe Especializada é aquela que possui atributos específicos, com mais detalhes – **SUBCLASSES** ou CLASSE DERIVADA.

As subclasses apontam para uma classe mais geral – as **SUPERCLASSES** ou CLASSE-BASE.

A HERANÇA é a capacidade de que uma subclasse herde todos os atributos, operações e relacionamentos de uma superclasse.

MOO – Herança (Generalização e Especialização) [2]



A notação para Herança é uma

seta partindo das sub-classes para a super-classe.

MOO – **Herança** (Generalização e Especialização) [3]

- Mostra uma relação entre um elemento mais geral para um mais específico.

- O elemento mais específico agrega mais informação ao geral.

- Agregação e Herança aplicam-se a outros classificadores (Diagramas)
  - não somente classes,
  - mas também: Pacotes, Casos de Usos, ...

MOO – Herança (Generalização e Especialização) [4]

Quando usar Herança ?

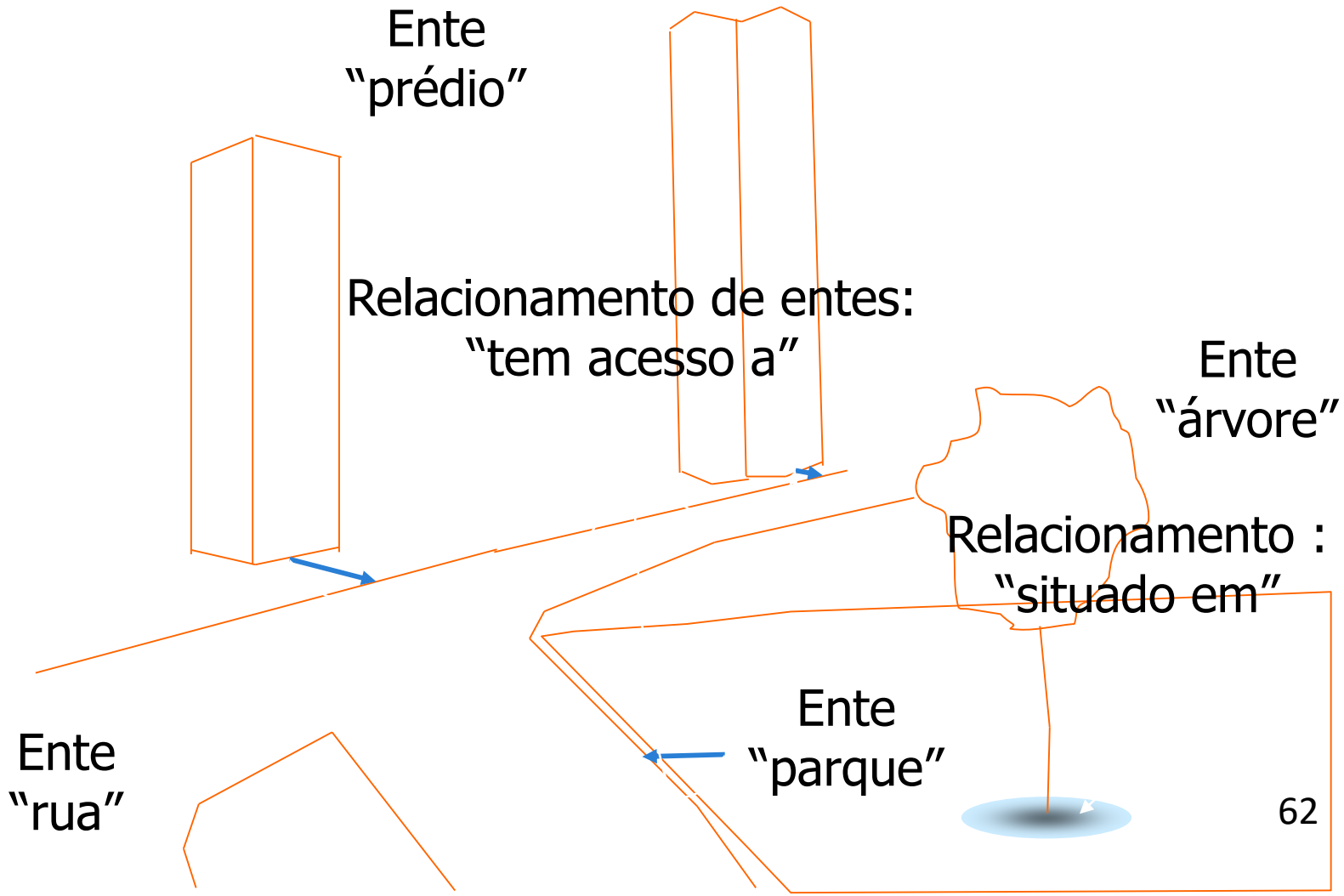
- Pergunta: A Classe filha é do tipo da mãe?
- Quando existe uma hierarquia entre Classes.
- Quando existe uma classificação (tipos).
- Para criar variações de uma classe.
- Frases:
  - A é do tipo de B
  - B é um A

- Para reaproveitar definições e simplificar o projeto.

Representação Computacional

Fonte: Carlos Vieira (Geocapacitar/UFV)

Representação Computacional: relacionamentos



Fonte: Carlos Vieira (Geocapacitar/UFV)

UML - Unified Modeling Language (1)

É uma notação que combina as 3 principais técnicas de Modelagem Orientada a Objeto: Modelagem OMT [Rumbaught *et al*., 1972], Análise e Projeto OO [Booch, 1994] e Objectory [Jacobon *et al*.,1992].

Adotada como notação padrão pela OMG (Object Management Group) como metodologia de modelagem de objetos.

É utilizada para descrever qualquer tipo de projeto de sistema por meio de diagramas, ex. **Diagramas de Classes, Casos de Usos**, não apenas em projetos de desenvolvimento de software, mas em projetos de implementações de sistemas de informações.

UML - Unified Modeling Language (2)

PRINCIPAIS CARCTERÍSTICAS:

Considera **Classificadores** na modelagem, que desempenham algum tipo de papel no projeto e na modelagem;

Os Classificadores podem ser: **Pacotes**, **Subsistemas**, Interfaces, Banco de Dados e **Classes**;

Os Relacionamentos de associações, herança e agregação desempenham o papel de relacionar os classificadores;

A partir de Pacotes e Subsistemas e demais classificadores e casos transacionais pode-se definir Casos de Usos e assim construir o modelo que realiza as transações especificadas por seus requisitos.

Questões para estudo

LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2011, 539p

Capítulo 8: Modelagem de Dados Geográficos [TÉCNICAS]

Pág 228

Questões para estudo (Capítulo 8)

Questão 2 - Por quê é útil incluir os níveis conceitual, lógico e de implementação na modelagem de dados geográficos?

Questão 4 (modificada) - Explique, com exemplos (geográficos), porque os conceitos de objeto, classe, agregação e herança (especialização / generalização) simplificam a modelagem de sistemas geográficos. Compare essa modelagem (baseada em objetos) com os modelos de dados centrados na geometria, que modelam o mundo como pontos, linhas, áreas, TIN ou raster/células matriciais.

Questões para estudo (Capítulo 8)

Questão 1 (modificada) - A Figura 8.21 é uma foto aérea obliqua de parte da cidade de Kfar-Saba, em Israel. Liste as classes de objetos (incluindo seus atributos e comportamentos) e mostre num Diagrama UML (*Unified Modeling Language*) os relacionamentos entre classes que seriam apropriados para o estudo de um SIG desta cidade.

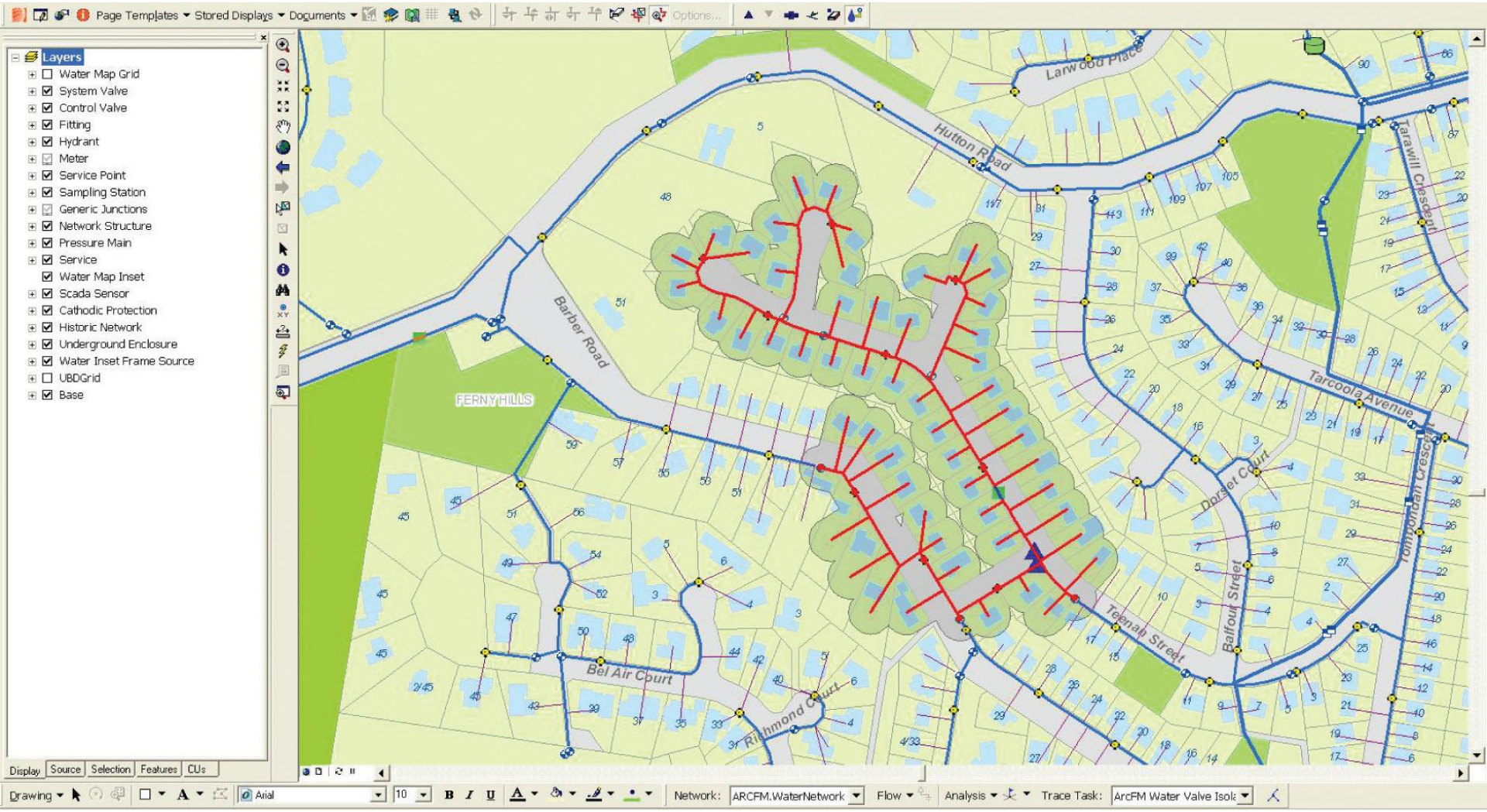
Questões para estudo (Capítulo 8)

Figura 8.21 Vista aérea oblíqua de Kfar-Saba, Israel. (Cortesia Av-revaya Mordehay)

Exemplo de um Modelo de Dados de Objeto de uma Rede de Distribuição de Água

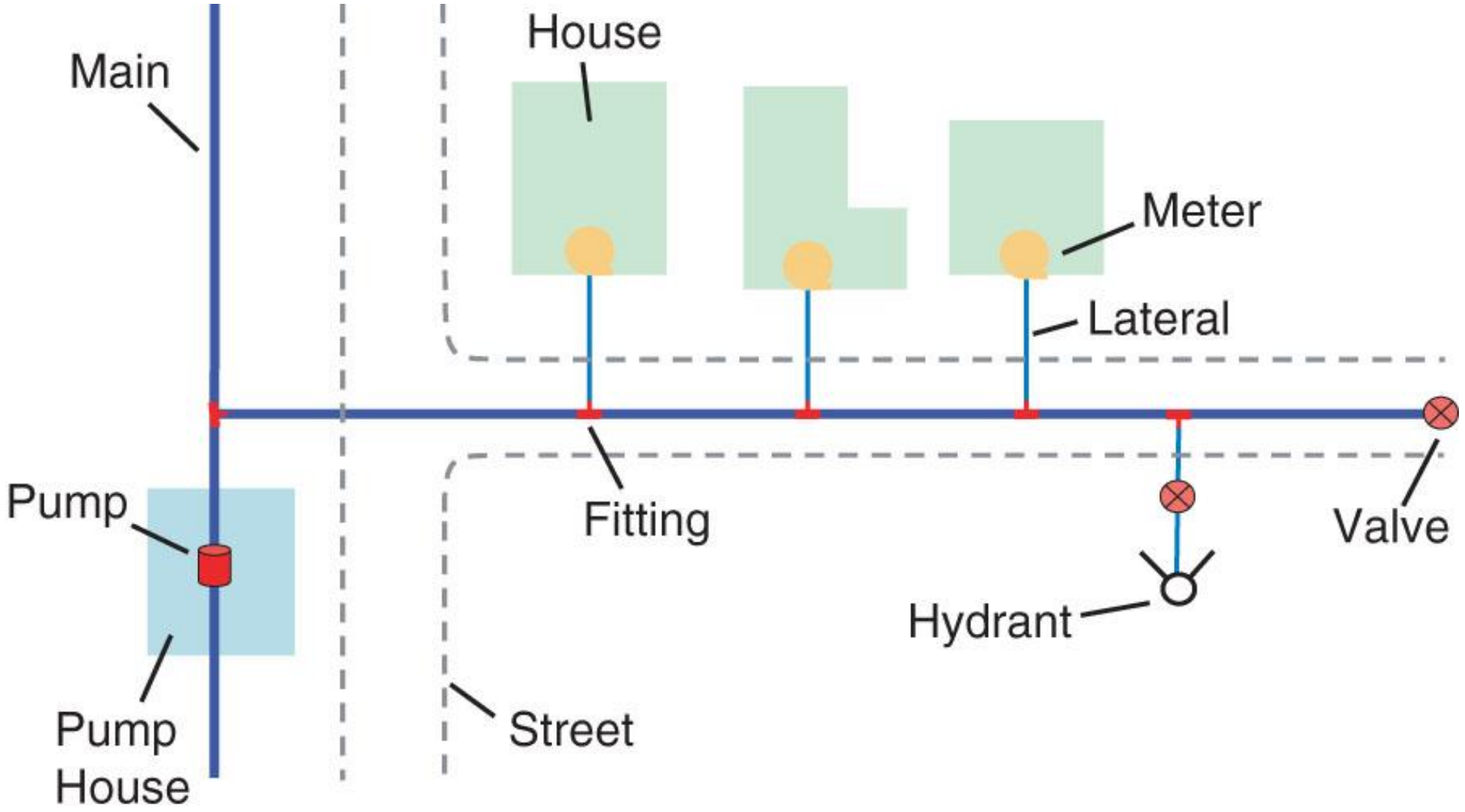
Seção 8.3: Figuras 8.16, 8.17, 8.18 e 8.19

“Utilities”: Água/Esgoto

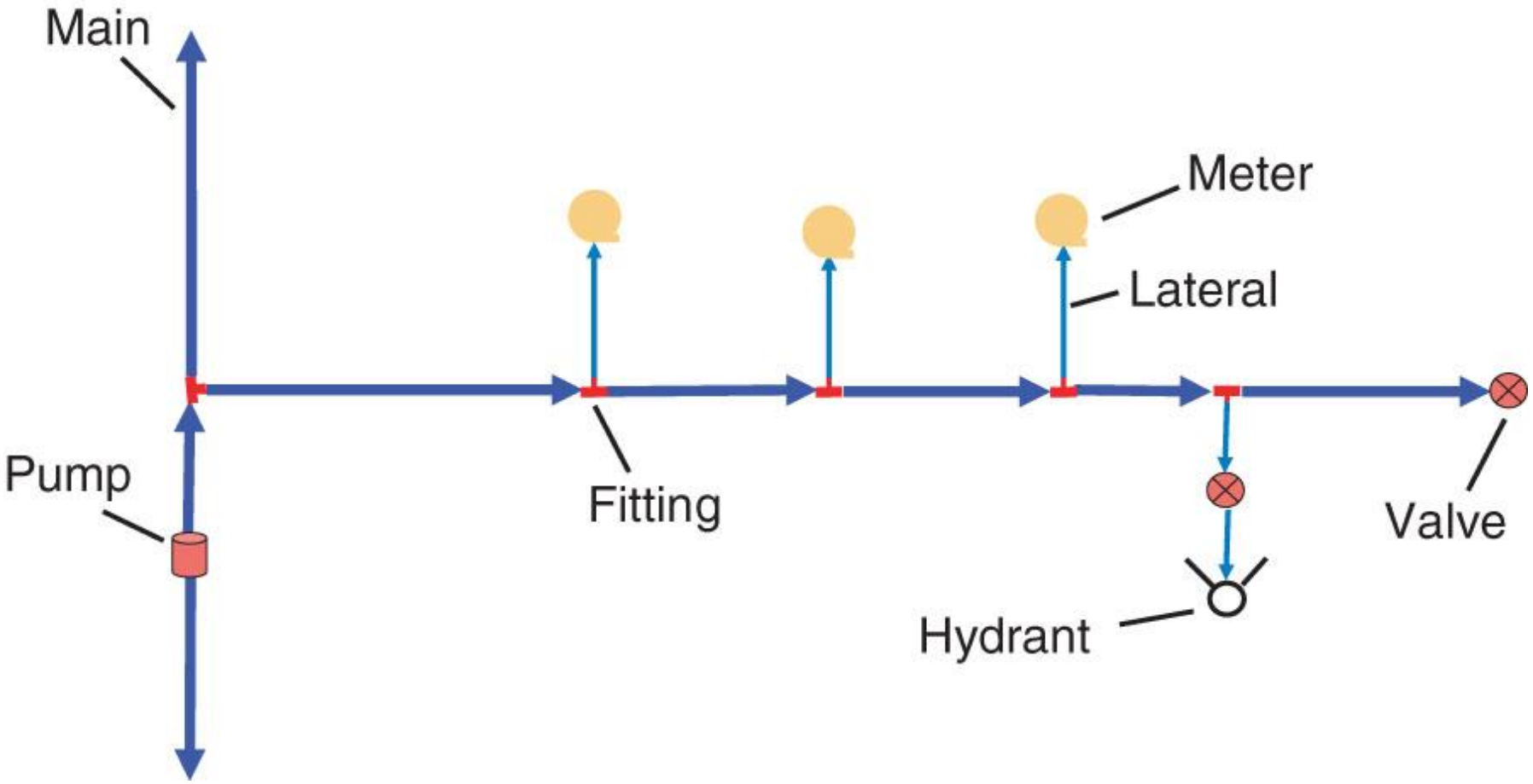


Fonte: Longley et al., 2013

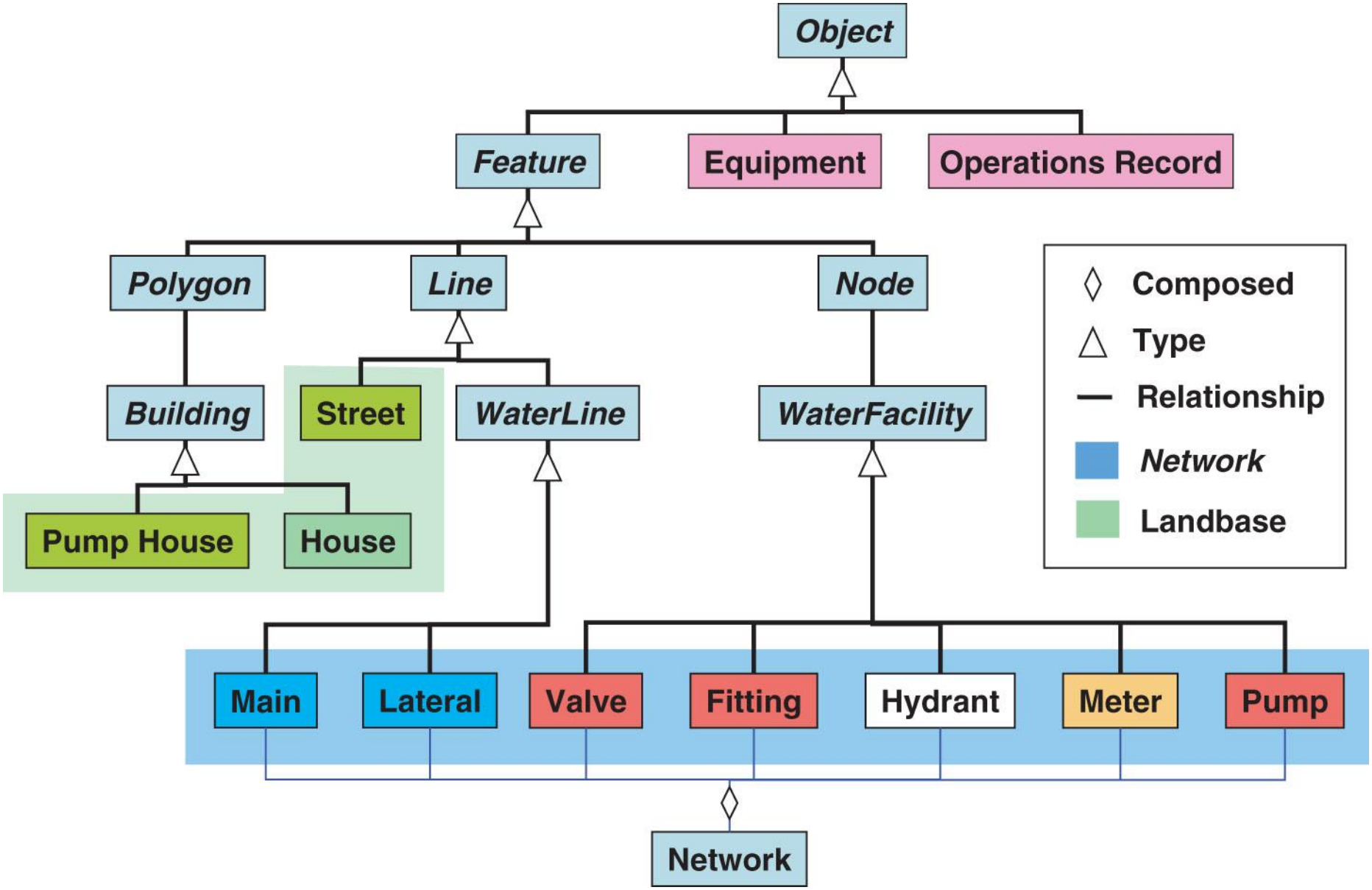
Water distribution system water-facility object types and geographic relationships



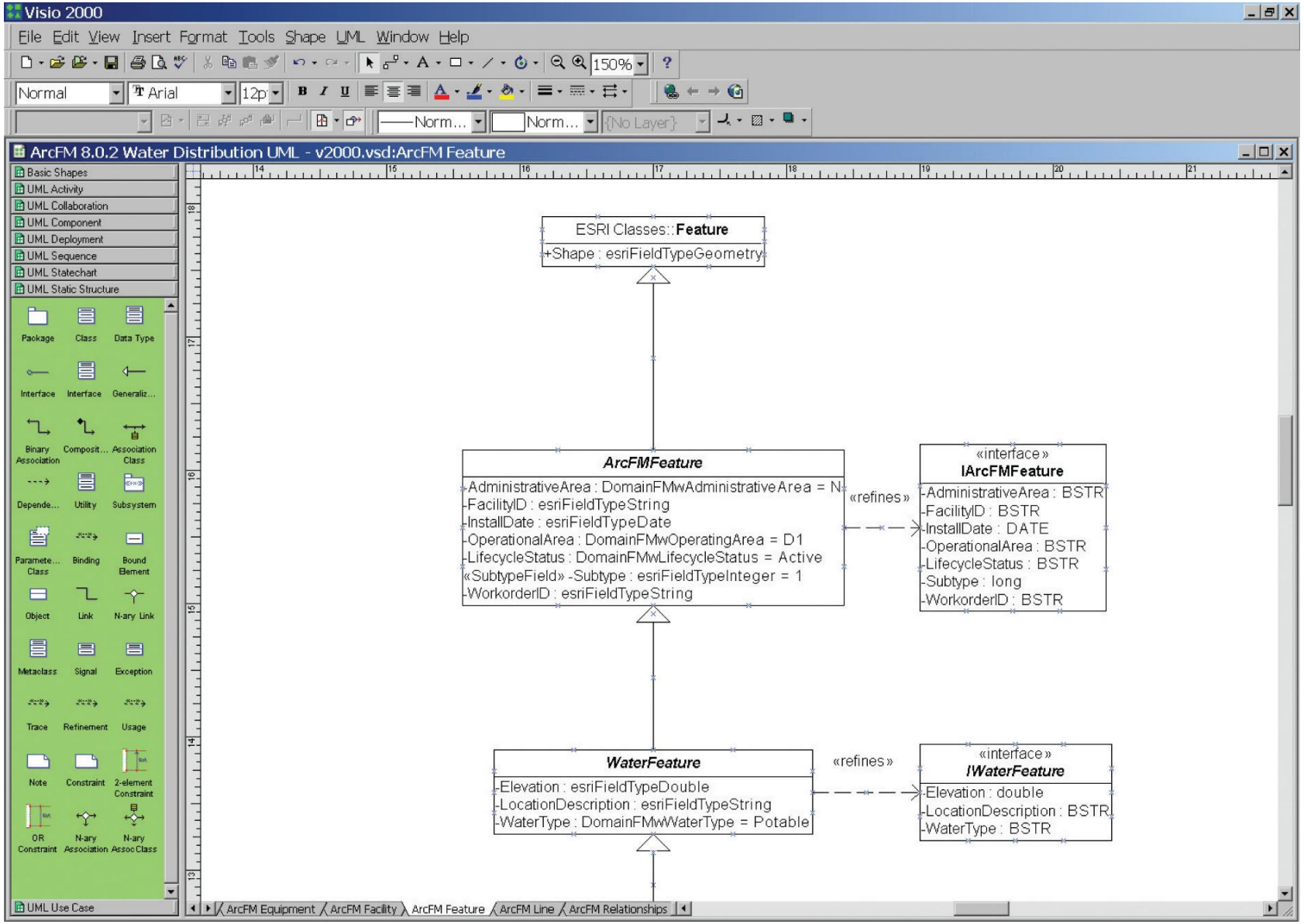
Water distribution system network



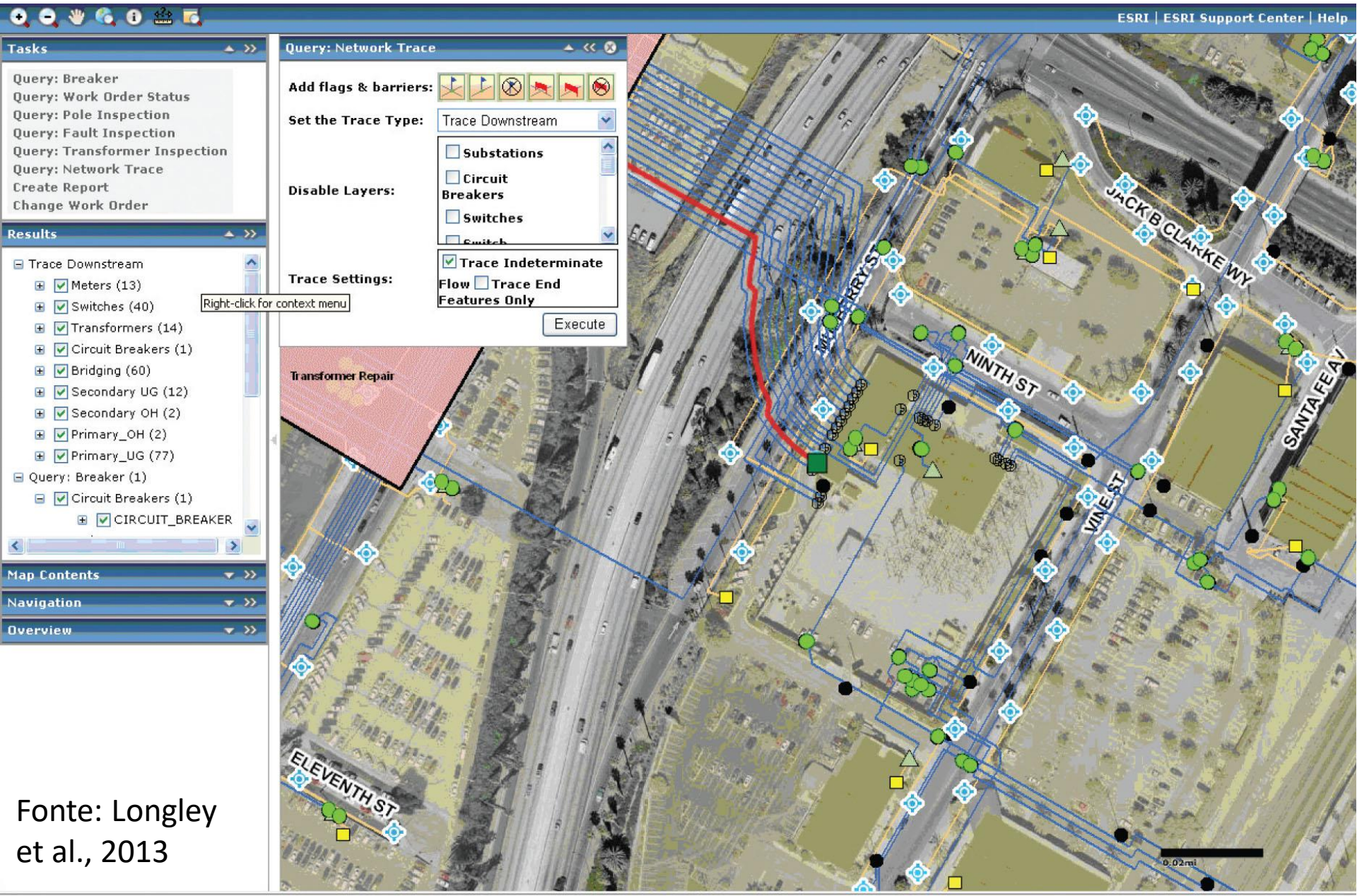
A water-facility object model



An example of a CASE tool (Microsoft Visio)
The UML model is for a utility water system

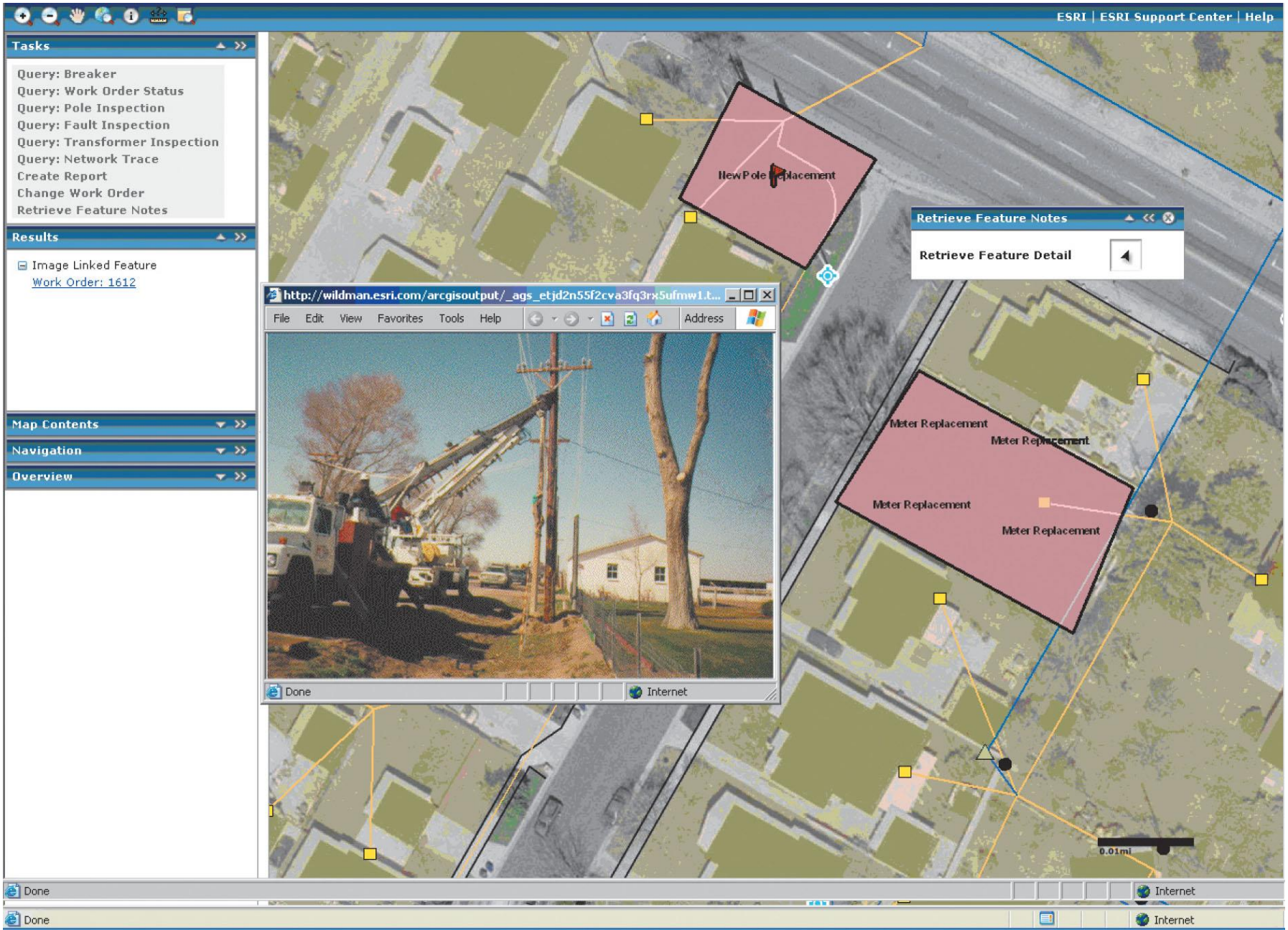


“Utilities”: Energia



Fonte: Longley et al., 2013

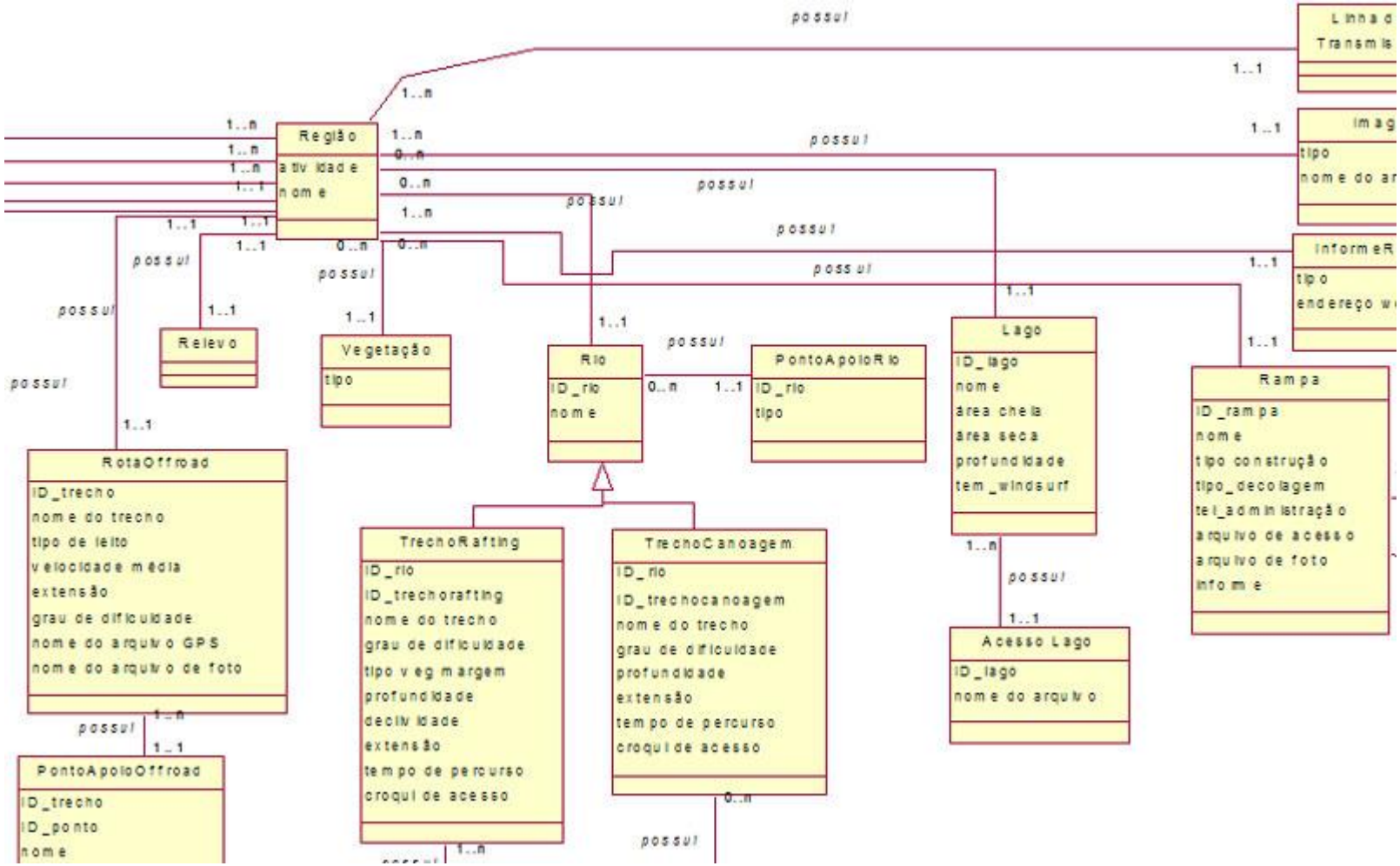
A photographic image used as a building object attribute in an electric facility system



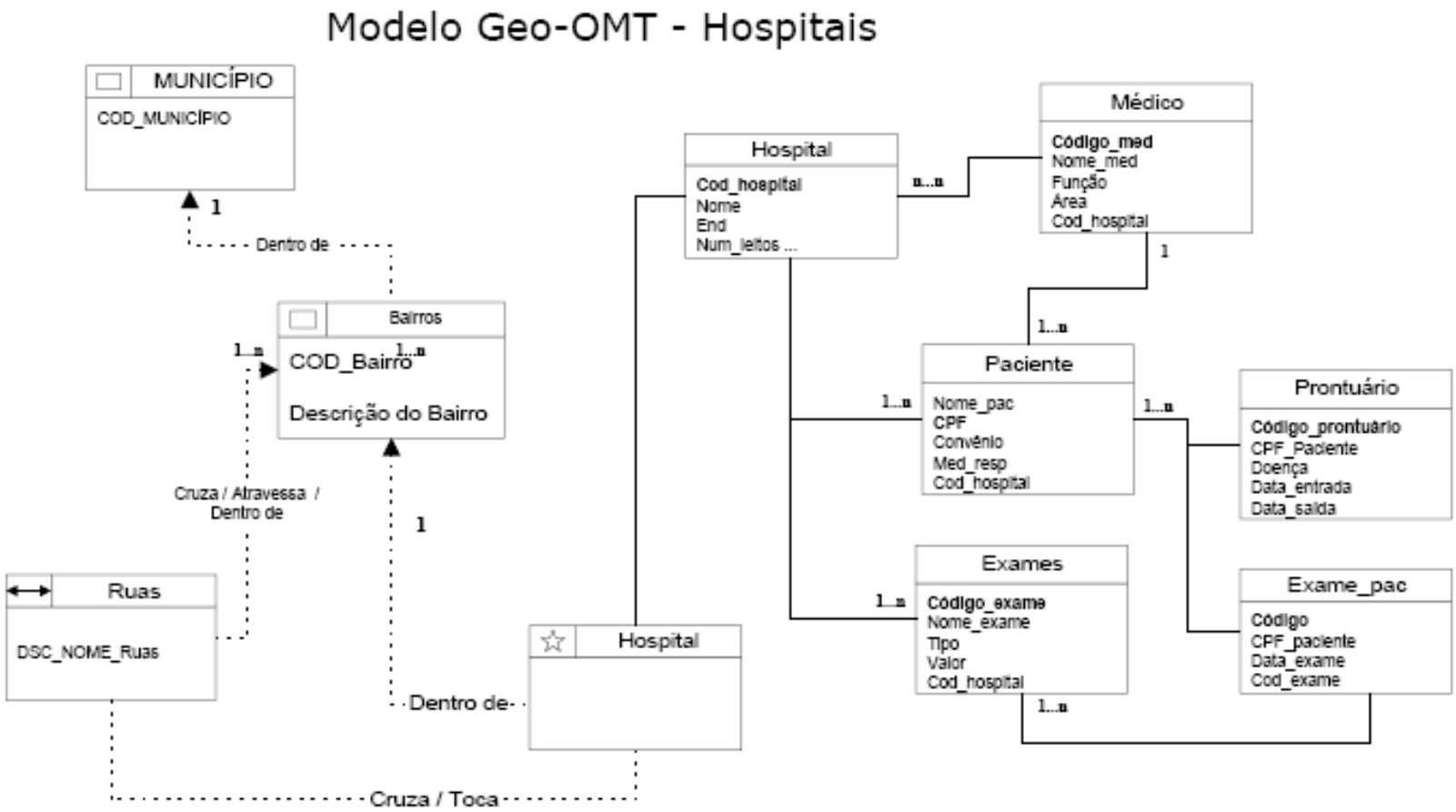
Técnicas de Modelagem de Dados Espaciais

Parte II – Dados Espaciais (OMT-G ou Geo_OMT)

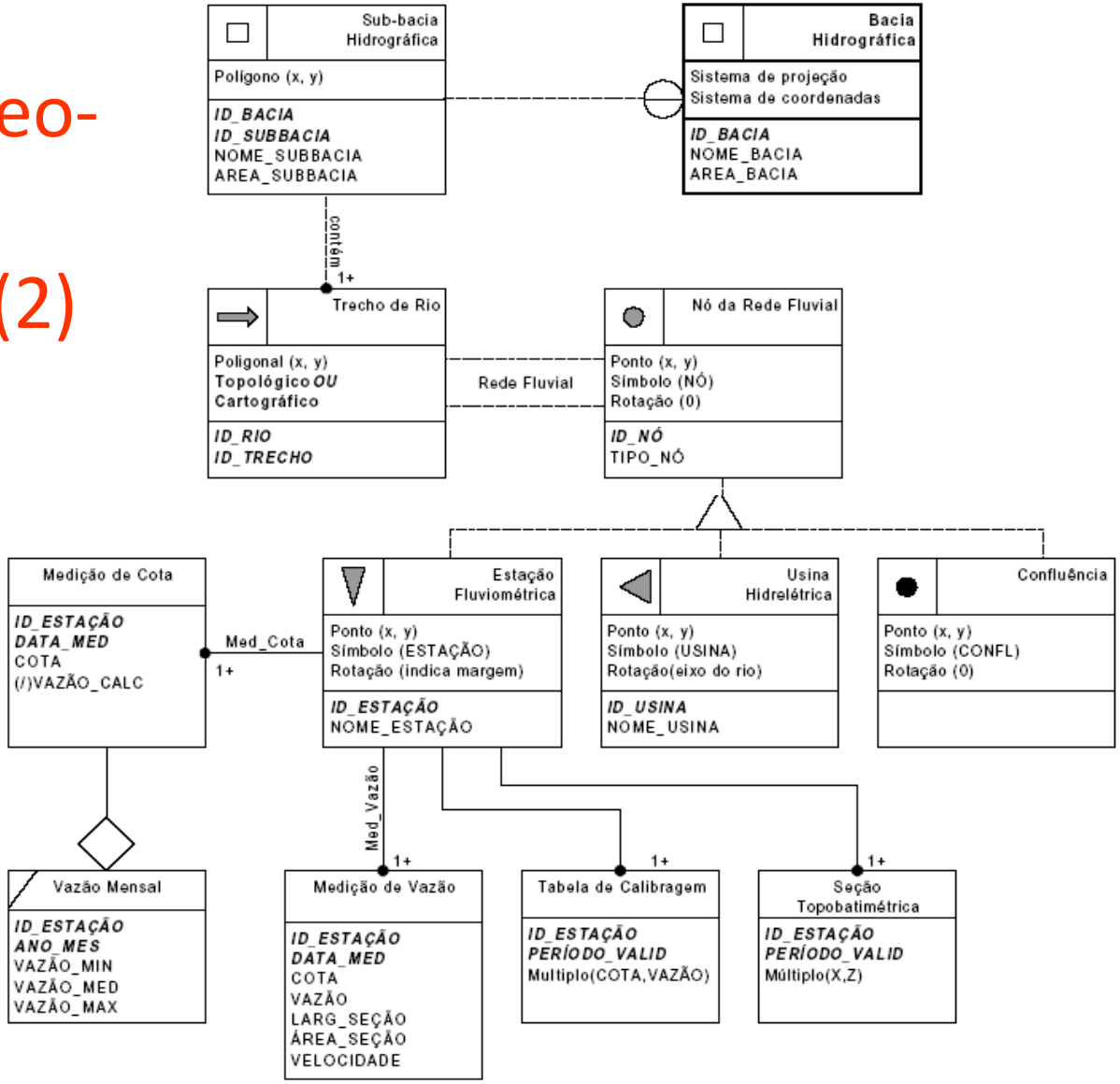
Modelagem da Serra da Mantiqueira



Modelo Geo-OMT – Exemplo (1)



Modelo Geo-OMT – Exemplo (2)



Coleta de Dados para SIG

Agenda 2

Bibliografia

- Modelagem de dados geográficos:

- LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2011, 539p

  Capítulo 9: Coleta de Dados para SIG [TÉCNICAS]

- TEIXEIRA, A. L. A.; CHRISTOFOLETTI, A. **Sistemas de informação geográfica (dicionário ilustrado)**. São Paulo: Hucitec, 1997, 244p.

Sumário

1. Introdução
2. Fluxo de trabalho da coleta de dados
3. Captura de dados geográficos primários
4. Captura de dados geográficos secundários
5. Obtenção de dados de fontes externas (transferência de dados)

1. Introdução

- Definições
  - Coleta de dados (*Data gathering*)
    - Procedimento de levantamento, identificação e obtenção dos dados a serem inseridos em um sistema.

      (Teixeira e Christofoletti, 1997)

1. Introdução

- Coleta de dados
  - É uma das tarefas de SIG mais caras e demoradas, porém, importante
  - Há muitas fontes de dados geográficos
  - Há muitos métodos para se inserir dados num SIG.
  - Podem ser classificados em 2 tipos:
    - **Captura**
    - **Transferência**

LONGLEY et al. (2013), pg. 229

1. Introdução

- Captura de dados
  - É importante distinguir entre:
    - Captura de dados primários (medida direta): imagens de satélites
    - Captura de dados secundários (origem de outras fontes): mapas
    - Distinguir entre dados vetoriais e matriciais
- Transferência de dados
  - Envolve a importação de dados digitais de outras fontes (3º), via internet

LONGLEY et al. (2013), pg. 229

1. Introdução

- Outros termos utilizados para coleta de dados, segundo LONGLEY et al. (2013), pg. 230:
  - Captura de dados
  - Automação de dados
  - Conversão de dados
  - Transferência de dados
  - Tradução de dados
  - Digitalização

Embora tenham diferenças sutis, descrevem a mesma coisa:

Adição de dados geográficos a um banco de dados

LONGLEY et al. (2013), pg. 229

1. Introdução

Classificação de dados geográficos para fins de coleta, com exemplos de cada tipo

	Matricial	Vetorial
Primários	-Imagens de satélites de Sensoriamento Remoto -Fotografias aéreas digitais	-Medições de GPS -Medições de levantamentos topográficos
Secundários	-Mapas ou fotografias escaneadas -Modelos digitais de altitude a partir de curvas de nível e mapas topográficos	-Mapas topográficos -Banco de dados de topônimos (nome de lugares)

LONGLEY et al. (2013), pg. 230

1. Introdução

- O investimento em coleta e transferência de dados:
  - Pode atingir de 60 a 85 % do total do investimento da implantação de um SIG. (LONGLEY et al. (2013), pg. 231)
    - Isto pode ocorrer se os custos de pessoal forem excluídos

      ou
    - Isto vale para organizações, como prefeituras ou empresa de serviços públicos de porte médio e grande, que ainda não tem uma base de dados adequada
  - A manutenção dos dados (após a implantação) poderá ser mais complexa e mais cara ainda.

1. Introdução

Decomposição de custos estimados pelos autores para 2 SIG do tipo cliente-servidor típicos (em US$ 1.000)

Dados: 15 a 50% do custo

	10 Estações de trabalho		100 Estações de Trabalho	
	$	%	$	%
Hardware	30	3,4	250	8,8
Software	25	2.8	150	5,3
Dados	400	44,7	450	15,8
Equipe	440	49,2	2.000	70,2
Total	895	100,0	2.850	100,0

LONGLEY et al. (2013), pg. 231

1. Introdução

- O sucesso de um projeto em SIG está diretamente relacionado à escolha dos dados de entrada e ao adequado processamento dos dados visando um produto final de qualidade.

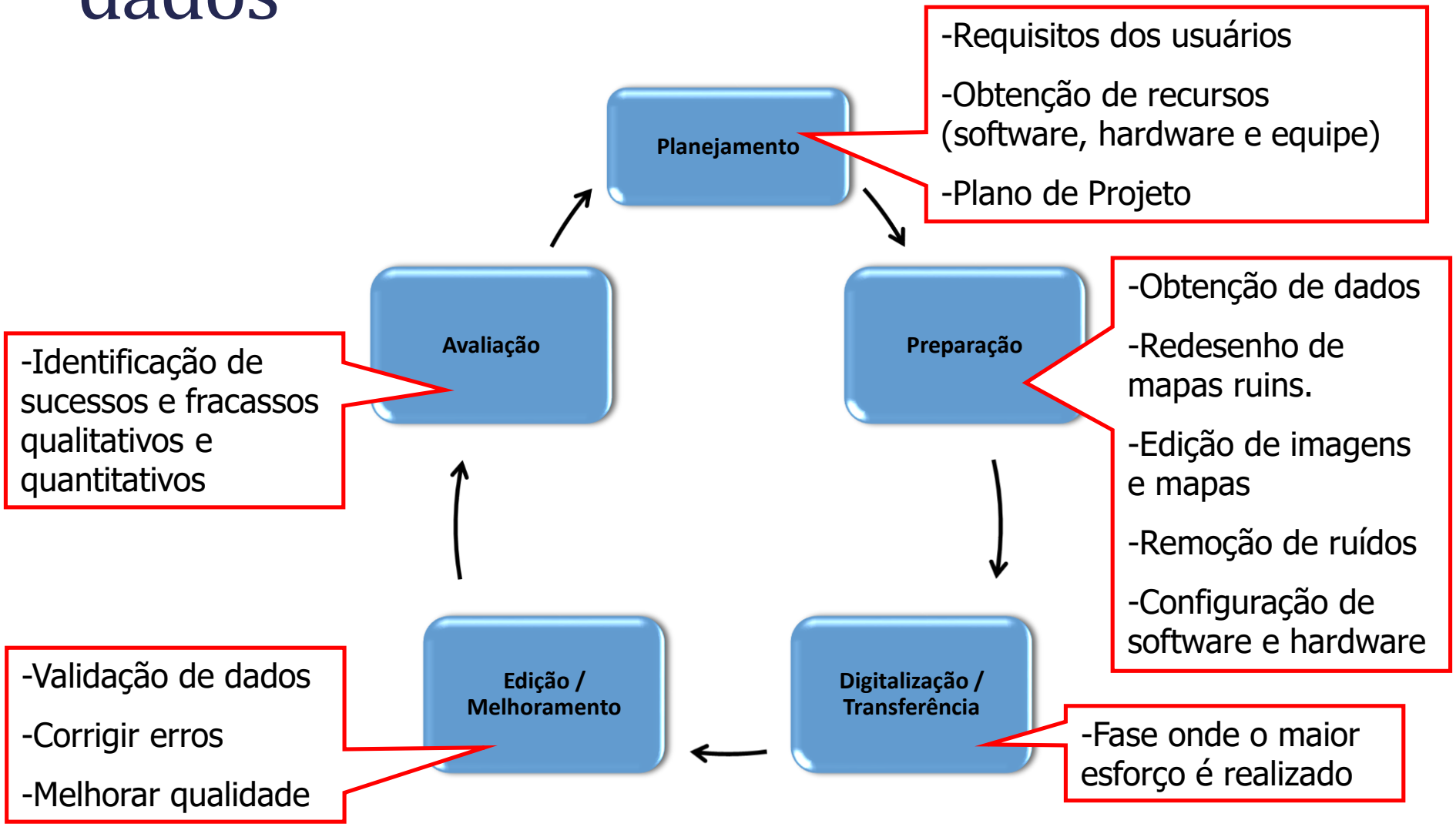
1. Introdução

- O processo de coleta e transferência de dados é um dos procedimentos mais importantes na implantação de um projeto utilizando SIG.

- Erros nesta fase podem atrasar ou mesmo inviabilizar a implantação de um projeto.

- Por isso, esta é considerada uma das fases mais caras, demoradas e tediosas de um projeto e este grande investimento inicial irá refletir na qualidade do produto final.

2. Fluxo de trabalho da coleta de dados



LONGLEY et al. (2013), pg. 231

3. Captura de dados geográficos primários

- Características:
  - Envolve a medição direta de objetos
  - Dados podem ser inseridos diretamente no banco de dados de um SIG
  - Minimiza a quantidade de tempo gasto
  - Minimiza possibilidade de erros
  - Há métodos para captura de dados:
    - Matriciais (raster)
    - Vetoriais

LONGLEY et al. (2013), pg. 232

3.1 Captura de dados matriciais

- O Sensoriamento Remoto (SR) é a forma mais popular de captura de dados matriciais primários
- SR é a medição de propriedades físicas, químicas e biológicas de objetos sem contato físico com eles
- Há 3 características fundamentais das imagens em SR:
- Resolução espacial, espectral e temporal

LONGLEY et al. (2013), pg. 234

Representação Matricial - exemplo

Imagens de Satélite

Mosaico de Imagens de Satélite do Brasil

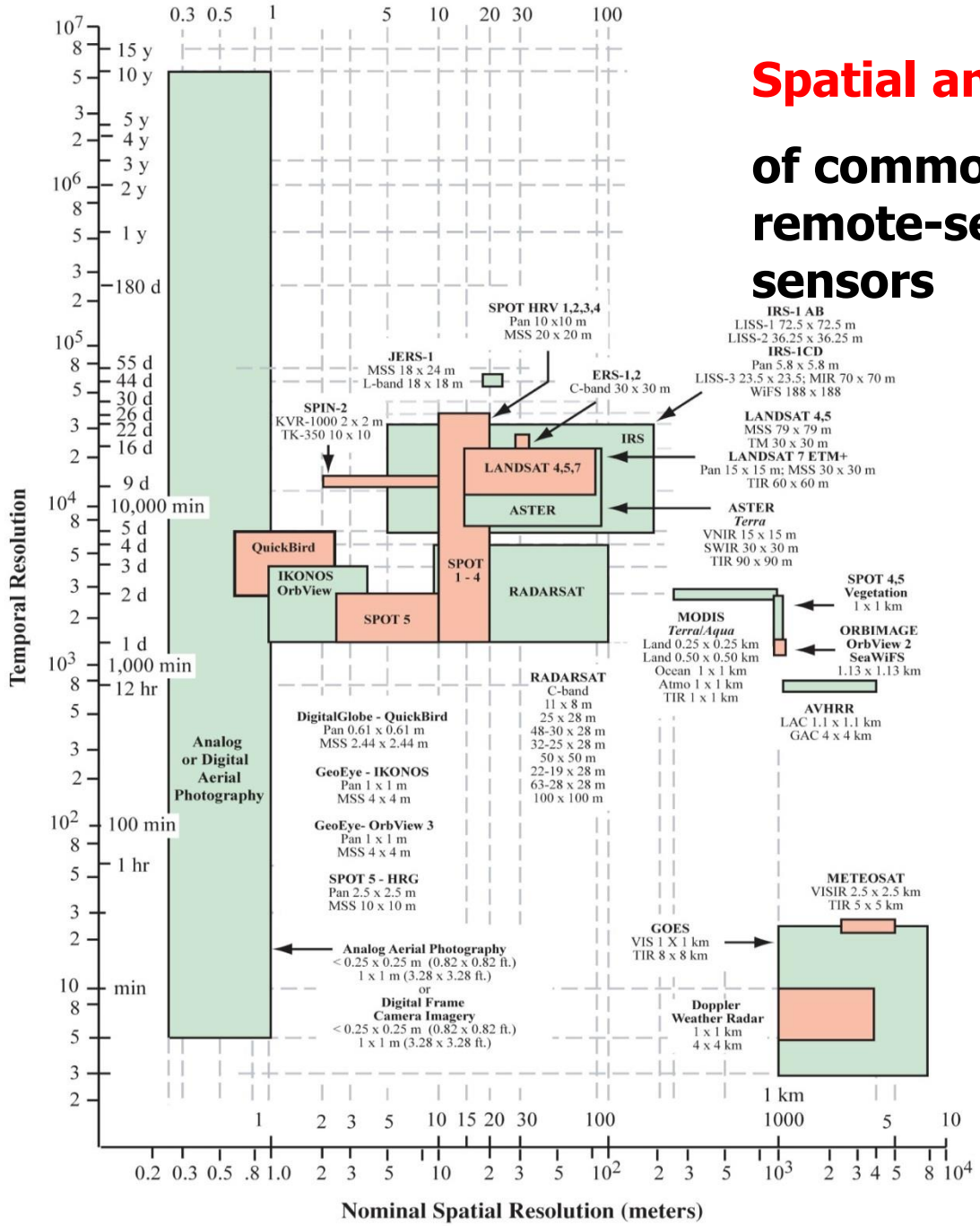
- Imagens aéreas digitais

imagens digitais obtidas a partir de câmera UltraCam.

Fonte:http://www.baseaerofoto.com.br/produtos/imagem-digital/

Spatial and temporal characteristics

of commonly used Earth observation remote-sensing systems and their sensors



(Source: John Jensen)

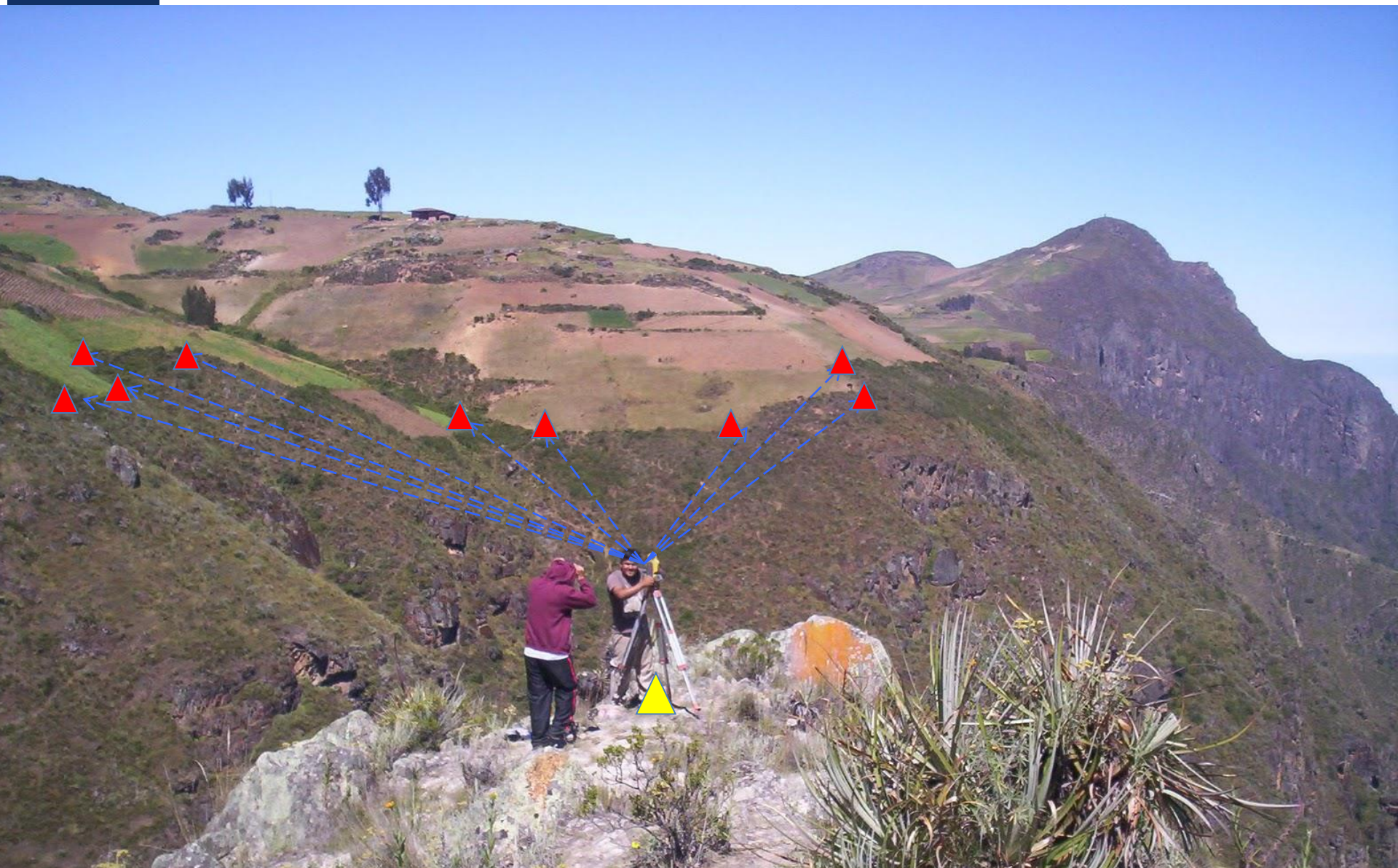
3.2 Captura de dados vetoriais

Principais formas de captura de dados

- Levantamentos Topográficos com Estação Total
  - Atividade muito demorada e cara
  - Melhor qualidade: alta acurácia

Fonte:http://www.dgyeatman.co.uk/

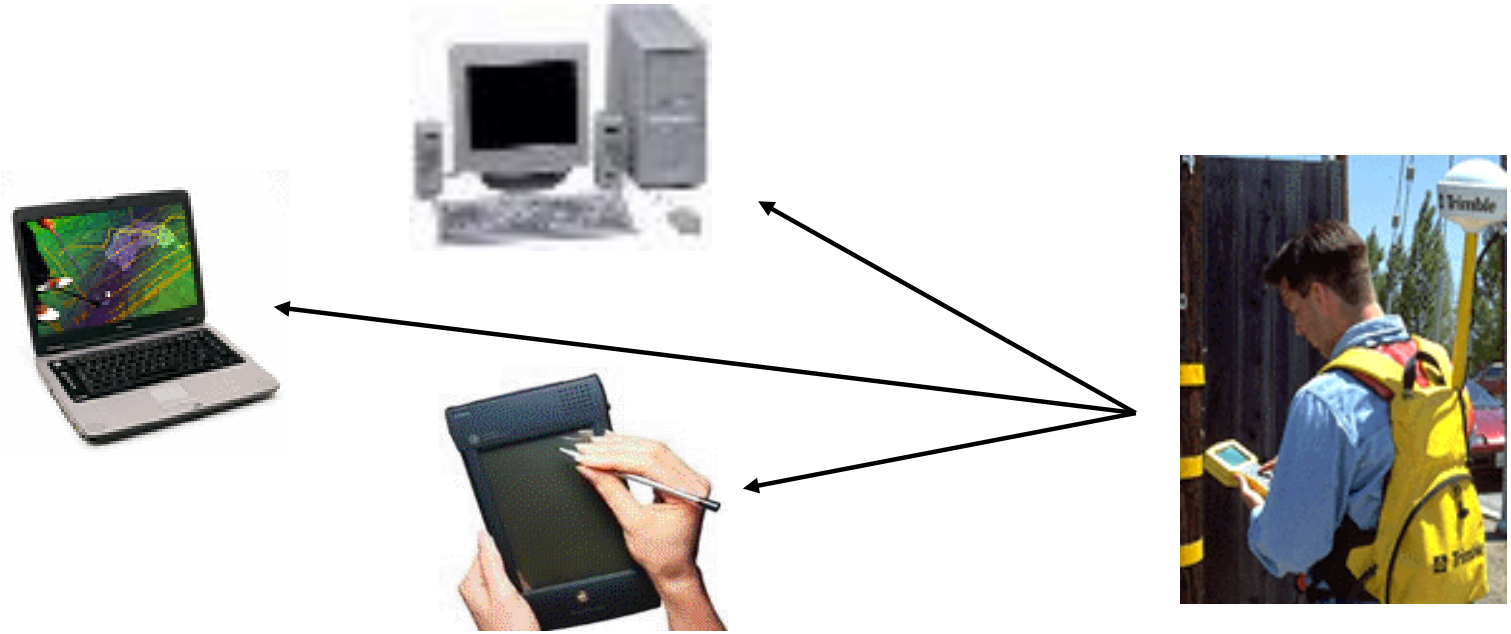
Coleta de Dados



3.2 Captura de dados vetoriais

- Dados coletados por GPS (Global Positioning System)

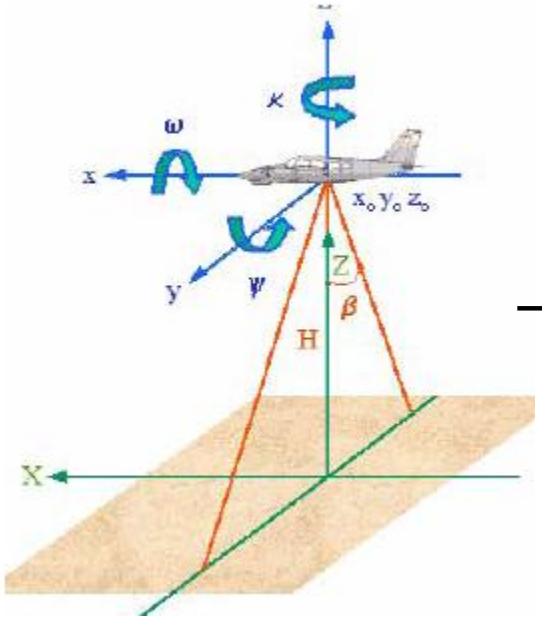
Os dados coletados em campo são importados para o SIG em formato digital, por meio de "palm tops", computadores portáteis ou diretamente nos computadores das estações de trabalho.

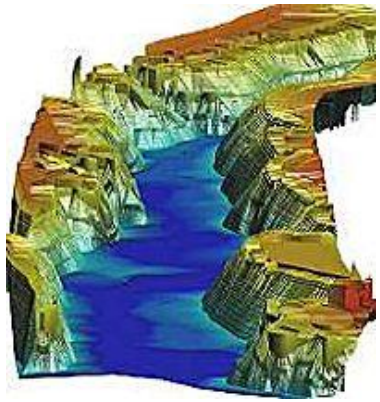


3.2 Captura de dados vetoriais

- Dados coletados por Laser Scanning.

Os pontos coletados durante o vôo são importados em formato digital para o SIG.

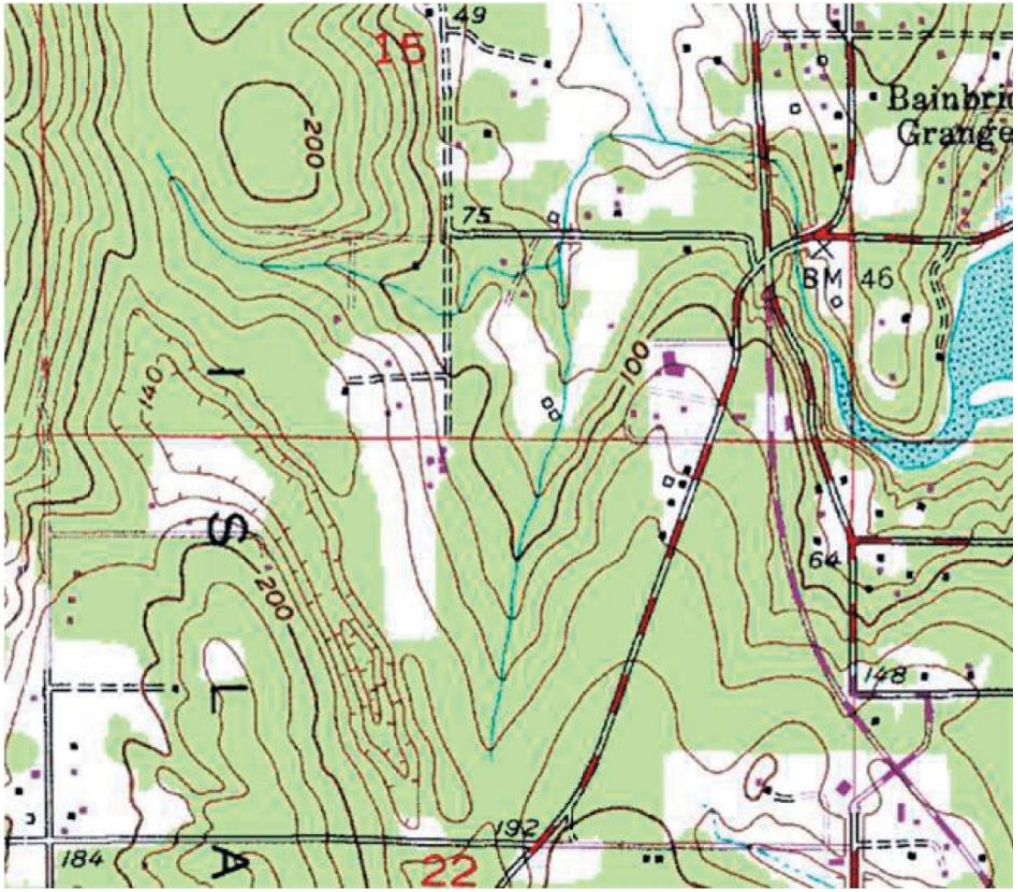




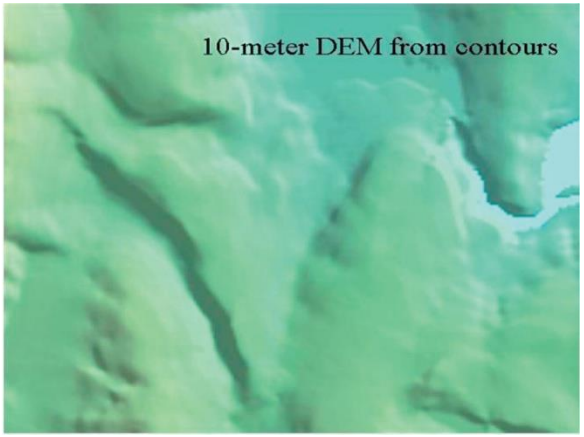
Comparison of three datasets for 1 square mile of Bainbridge Island, Washington State

(From www.pugetsoundlidar.ess.washington.edu/About_LIDAR.htm)

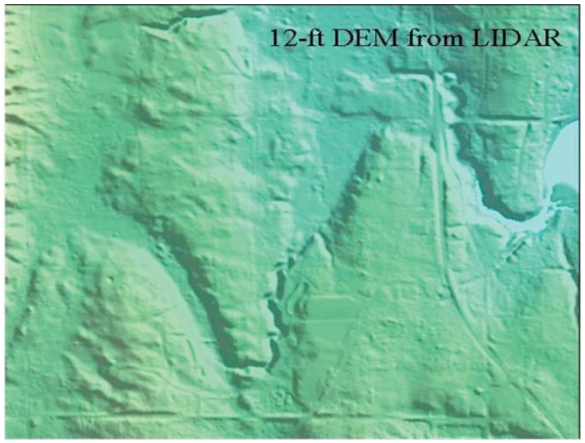
(A)



(B)



(C)



(A) scanned USGS 1:24,000 topographic map sheet

(B) 10 m digital elevation model (DEM) derived from contours digitized from a map sheet; (C) 12 ft (365-cm) resolution DEM derived from a LiDAR survey

4. Captura de dados geográficos secundários

- Características:
  - Processo de criação de arquivos matriciais e vetoriais de banco de dados a partir de:
    - Mapas
    - Fotografias
    - Outros documentos impressos
  - Há métodos para captura de dados:
    - Matriciais (escaneamento)
    - Vetoriais (digitalização em tela, fotogrametria estereoscópica, ...

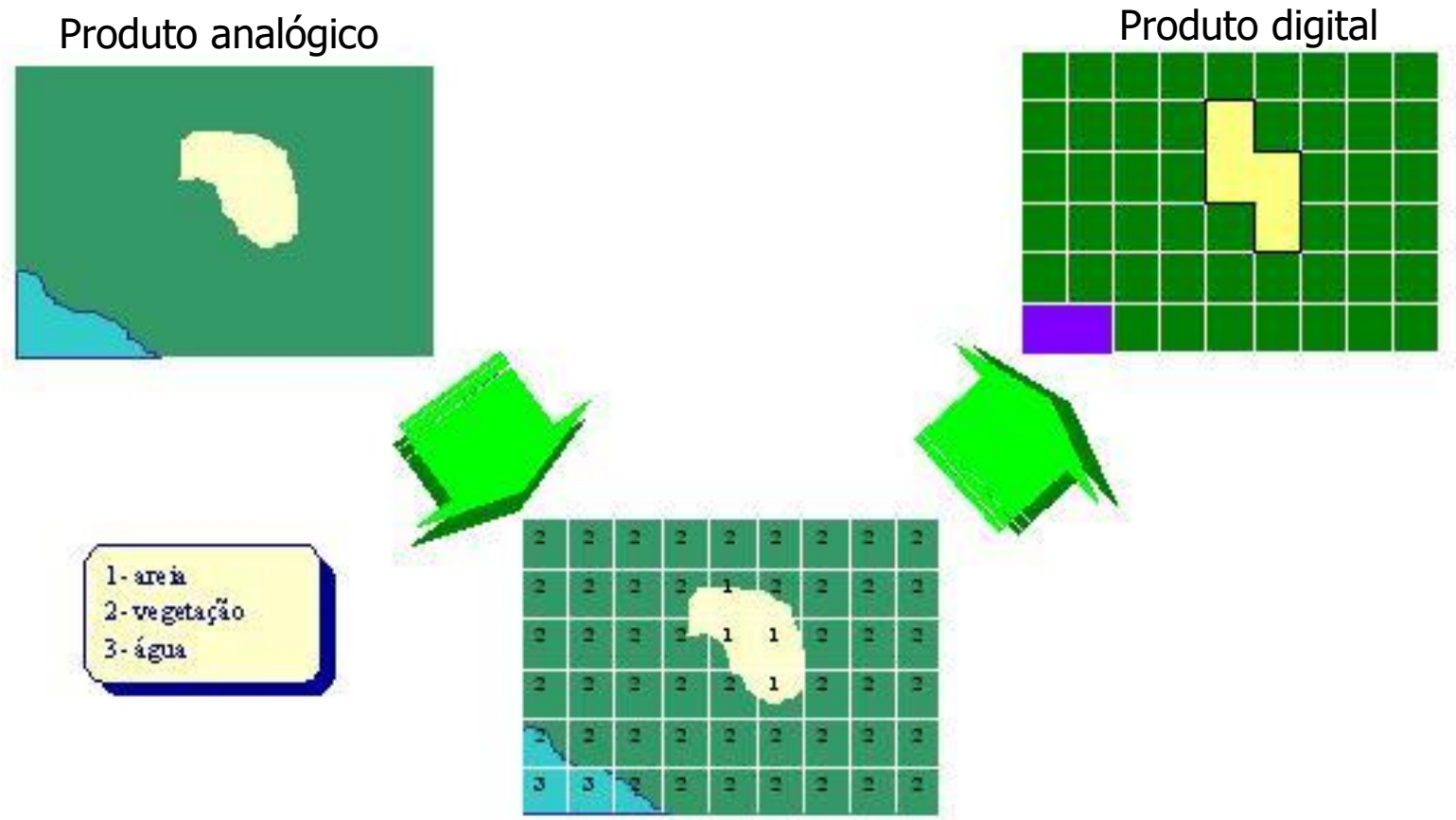
LONGLEY et al. (2013), pg. 236

4.1 Captura de dados matriciais

Entrada de Dados via Scanner

- É a digitalização de documentos analógicos por meio da conversão, em uma única etapa, utilizando um equipamento scanner.
- A qualidade dos produtos gerados depende tanto do equipamento quanto da complexidade do documento de entrada.
- Procedimento simples que consiste na passagem e leitura do documento pelo scanner, com a vantagem de ser realizado rapidamente, desde que os documentos já tenham sido classificados e o equipamento configurado.

Escanerização



Modelos de Scanner

- Scanner de mesa

- Scanner de tambor (drum)

- Scanner de largo formato (large format feed)

Principais problemas no escaneamento (1)

- Distorção óptica quando usado scanner doméstico
- Informações indesejáveis, tal como anotações à mão
- Formato dos dados produzidos e entrada no software SIG
- Volume de edições para produção de dados convenientes para análise
- Seleção de tolerância apropriada para assegurar a codificação de dados de interesse e os dados com fundos indesejáveis

Principais problemas no escaneamento (2)

- Qualidade e complexidade do documento de entrada;
- Baixa resolução do scanner;
- Má discriminação entre os elementos gráficos do documento;
- Pouca memória na máquina;
- Falta de experiência do operador.

Custos envolvidos no processo de escaneamento

- Mais altos quanto mais complexo for o material de origem;
- Proporcional a quantidade de material;
- Processos de conversão após o uso do scanner;
- Aplicativos gráficos para edição e melhoramento da qualidade da imagem;
- Scanners com boa resolução são caros;
- Operador com boa experiência no manuseio da máquina é bem remunerado

4.2 Captura de dados vetoriais

4.2.1Conversão na tela do monitor “Heads up”

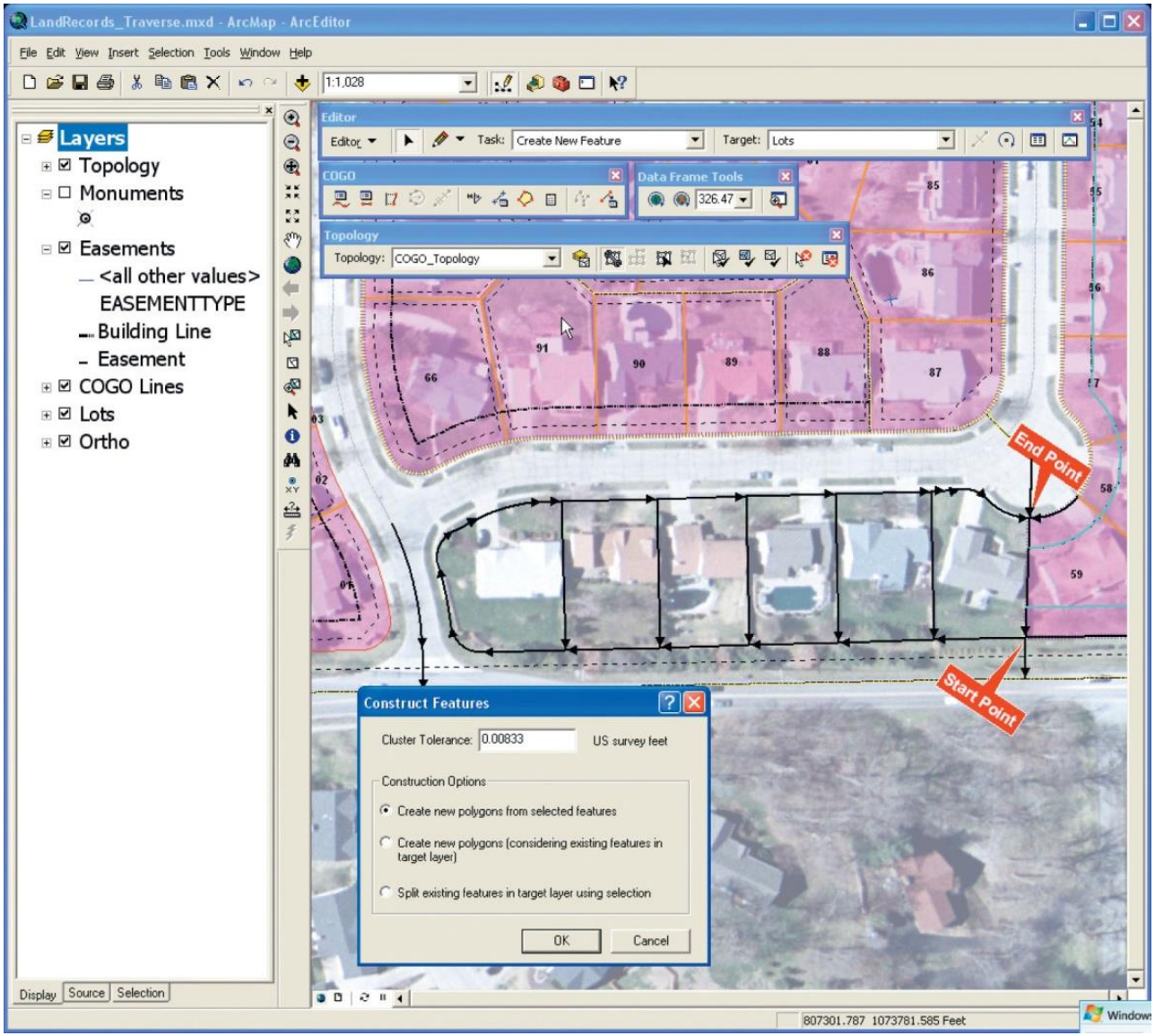
- É a extração de entidades (ponto, linha e área) diretamente da tela do computador utilizando uma imagem (raster) ao fundo como fonte de informações

- operador está posicionado de maneira a observar a tela de computador (com a cabeça para cima)

- A precisão da digitalização na tela é alta, pois os dados são escaneados em altas resoluções (200 a 1600 DPI) e a fadiga do operador é minimizada neste processo

- As ferramentas de “zoom” ajudam no processo de digitalização.

An example of raster background data (color aerial photography) underneath vector data (land parcels) that are being digitized on-screen



4.2 Captura de dados vetoriais

4.2.2. Conversão Semi-automática

- É o processo de conversão no qual ocorre intervenção do operador em etapas de informações duvidosas
- Identifica e reconhece bordas
  - Decide se é linha ou não e qual a continuidade da linha

- Necessária boa qualidade do dado a ser digitalizado
- Necessário processo de edição e correção de erros
- A qualidade do produto final dependerá do operador e do arquivo raster
- Utilização: desenhos extremamente complexos

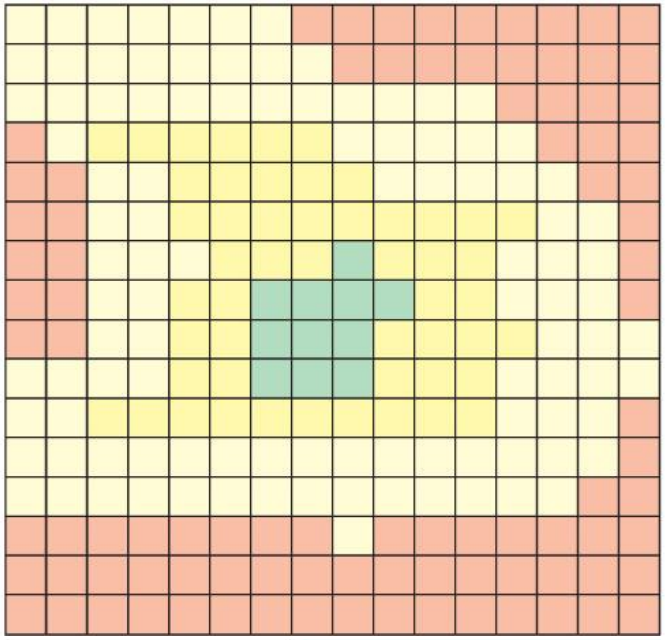
4.2 Captura de dados vetoriais

4.2.3. Conversão Automática

- É o processo de conversão matricial para vetorial ou vetorial para matricial, realizado através de aplicativos específicos de conversão automática sem intervenção de operador

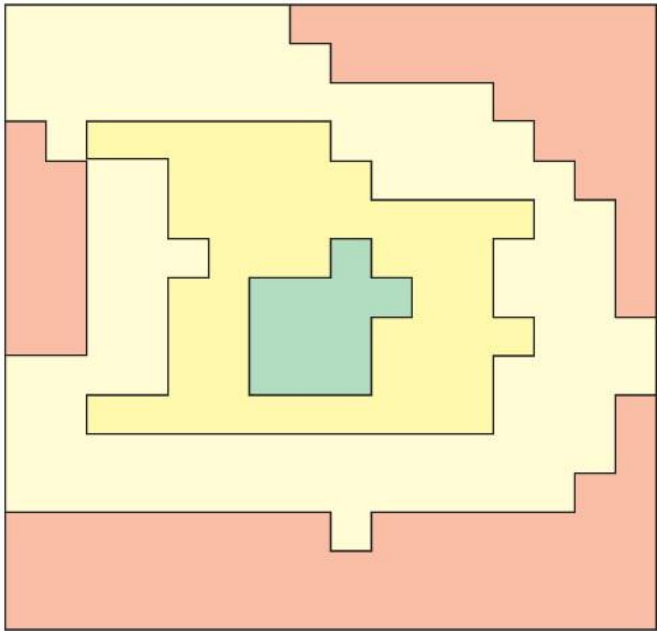
  - As linhas são criadas automaticamente através de reconhecimento de padrões;
  - Rapidez na criação dos arquivos vetoriais;
  - É necessário uma boa qualidade do dado a ser digitalizado;
  - É necessário a edição e a correção de erros.

Batch vectorization of a scanned map



A

(A) original raster file



B

(B) vectorized polygons

Adjacent raster cells with the same attribute values are aggregated. Class boundaries are then created at the intersection between adjacent classes in the form of vector lines

4.2 Captura de dados vetoriais
4.2.3 Entrada de dados tabulares

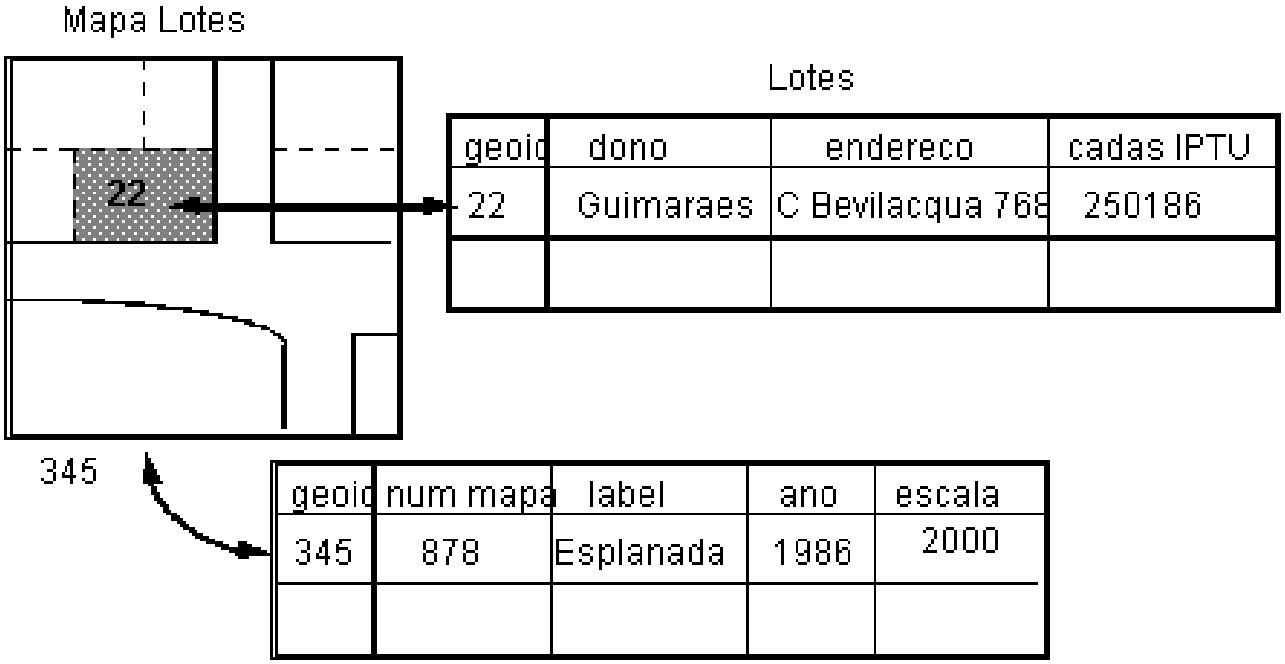
- Dados tabulares são informações descritivas das entidades espaciais representadas nas formas matricial ou vetorial, organizadas na forma de tabelas.

- A ligação entre um Sistema de Informação Geográfica (SIG) e um Banco de Dados Relacional se dá por meio de um Sistema Gerenciador de Banco de Dados Relacional - SGBDR

- Os componentes espacial e descritivo do objeto geográfico são armazenados separadamente – numa das arquiteturas.

4.2.3 Entrada de dados tabulares

Na Arquitetura Dual os atributos convencionais são guardados nas tabelas e os dados espaciais são armazenados em um sistema específico. A conexão é feita através de identificadores (id) de objetos.



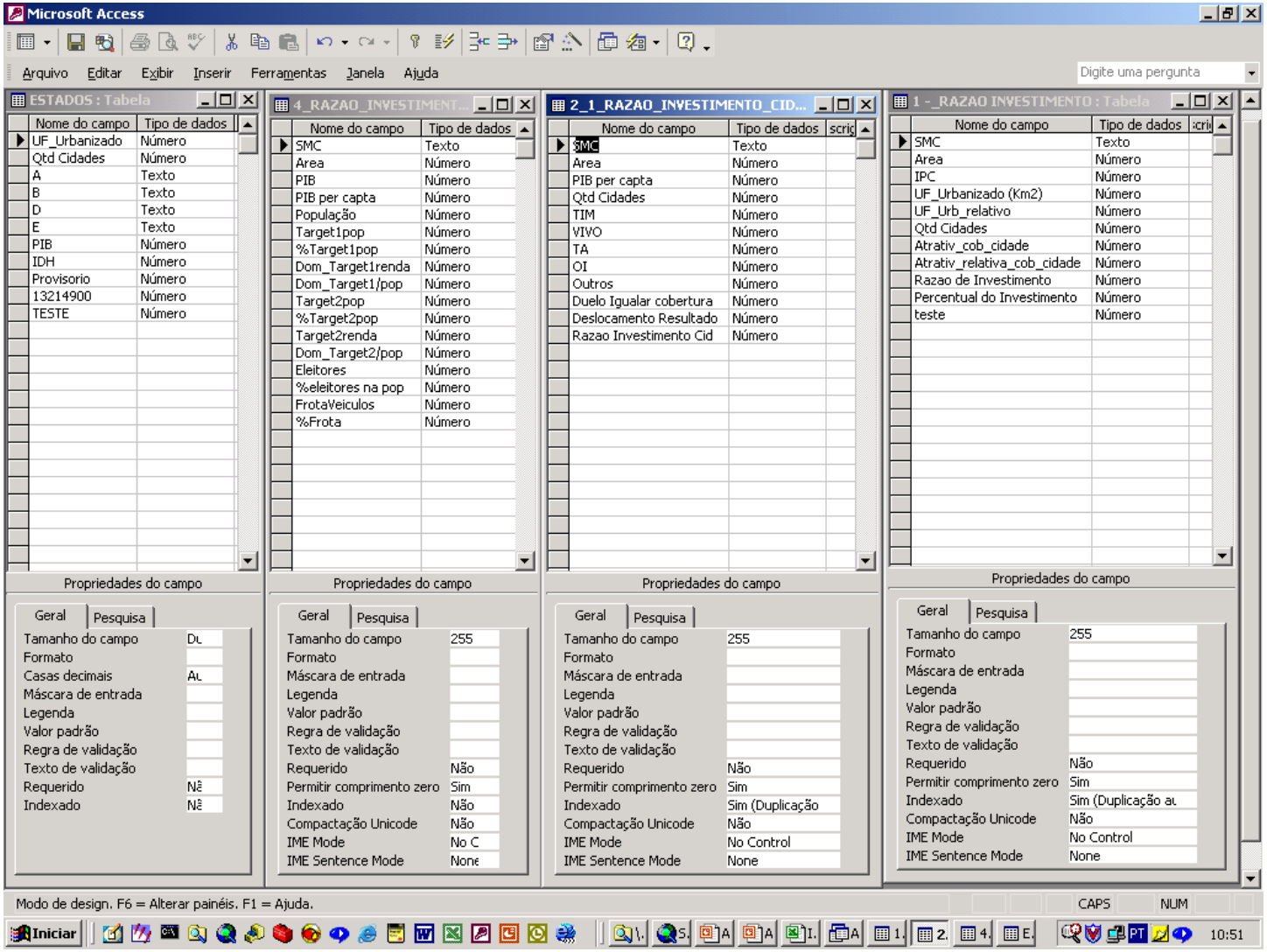
4.2.3 Dados Tabulares - exemplos

Attributes of SP_TAB1

X	Y	CNPJ	EMPRESA	UL	DVS	RAZAO	ENDERECO
5738509,186831	4057144,438	61749107000	61749107	0002	02	AGROPECUARIA RIO BRILHANTE LTDA	R DR ALCEU DE C RODRIGUES 229
5742257,939602	4069566,022	64958606000	64958606	0001	17	PARAISO DE SHALON PLANTAS E FLORES LTDA	RUA ORLANDO PELEGRINO 13
5734708,535691	4050351,117	03481074000	03481074	0001	85	AGRO PECUARIA TAPIRAPE SA	RUA AMADOR BUENO 000474 2C075
5743334,456083	4062780,198	03487501000	03487501	0002	13	AGROPECUARIA SAVANA S/A	AL DINO BUENO 715
5734708,535691	4050351,117	15026073000	15026073	0002	00	AGROPECUARIA RIO DARRO	RUA AMADOR BUENO 000474
5734708,535691	4050351,117	15026073000	15026073	0002	00	AGROPECUARIA RIO DARRO	RUA AMADOR BUENO 000474
5742776,466451	4060826,937	53441424000	53441424	0001	58	SANDEX EMPREENDIMENTOS AGROPASTORIL LTDA	PRACA DA LIBERDADE 000130 11 ANDAR CJ 111
5742776,466451	4060826,937	65647331000	65647331	0001	63	SANMAX AGROPASTORIL LTDA	PRACA DA LIBERDADE 000130 11 ANDAR CJ 111
5742776,466451	4060826,937	65647356000	65647356	0001	67	AGROPECUARIA AGUA LIMPA LTDA	PRACA DA LIBERDADE 000130 11 ANDAR CJ 111
5738509,186831	4057144,438	65836017000	65836017	0002	00	AGROPECUARIA TAMBARU LTDA	R DR ALCEU DE C RODRIGUES 229
5741799,792018	4057654,108	68075076000	68075076	0001	10	AILE COMERCIO E PARTICIPACOES LTDA	RUA DR. FABRICIO VAMPRE 000070
5745284,841621	4051726,190	00872437000	00872437	0001	15	AVICULTURA JO LTDA ME	AV PADRE ARLINDO VIEIRA 002302
5747492,360599	4072164,143	02116808000	02116808	0001	64	ÆROMA COMERCIO DE ANIMAIS EXOTICOS LTDA	RUA MATEUS CAMPOS 165
5734708,535691	4050351,117	01301035000	01301035	0002	04	AGROPECUARIA ALVORADA DO NORTE	RUA AMADOR BUENO, 474 0
5734708,535691	4050351,117	04814570000	04814570	0002	57	AGROPECUARIA TANGURO S.A	RUA AMADOR BUENO, 474 0
5737416,06107	4048077,953	17978388000	17978388	0002	83	GT AGRO CARBO INDUSTRIAL LTDA	RUA MONLEVADE 000325
5737416,06107	4048077,953	17978388000	17978388	0002	83	GT AGRO CARBO INDUSTRIAL LTDA	RUA MONLEVADE 000325
5742653,883667	4061331,189	60387784000	60387784	0001	84	AGROPECUARIA POUSO ALEGRE	AV. BRIGADEIRO LUIZ ANTONIO 000290 11 ANDAR
5737607,125646	4047282,723	00442510000	00442510	0001	19	CENTRAL DE ATENDIMENTO REPRESENTACAO S/C LTDA	AV. NOSSA SENHORA DO SABARA 004350
5736655,089933	4054829,963	02038748000	02038748	0001	09	MARIA AUGUSTA DOS SANTOS HORTI FRUTI GRANJEIROS ME	RUA JAMES WATT 000128
5741916,25653	4062769,469	59479006000	59479006	0001	36	DOMINGOS J P OLIVEIRA SC LTDA	AL BARAO DE LIMEIRA 000478
5734712,883851	4039895,949	61067187000	61067187	0001	35	JARDINS NATUREZA PAISAGISMO SC LTDA ME	AV SENADOR TEOTONIO VILELA 006530
5738316,929807	4056694,665	65686206000	65686206	0001	62	EDISON ALBINO SILVA ME	RUA JOSE GONCALVES PEREIRA 000105
5742328,952594	4061908,492	66723586000	66723586	0001	20	FORMOSA S/A AGRICULTURA IND. E COMERCIO.	RUA BARAO DE ITAPETININGA, 000140 15 ANDAR
5739262,543503	4057552,457	71977623000	71977623	0001	67	NUCLEO VERDE PAISAGISMO LTDA-ME	RUA TAJA 211
5742776,466451	4060826,937	03825056000	03825056	0001	73	SANYO AGROPECUARIA LTDA	PRACA DA LIBERDADE 000130
5734708,535691	4050351,117	01071445000	01071445	0001	25	VALPINUS AGROINDUSTRIA FLORESTAL LTDA	RUA AMADOR BUENO 000474
5742653,883667	4061331,189	58278367000	58278367	0001	51	SOCIEDADE AGRICOLA SANTA CAMILA LTDA	AV BRIGADEIRO LUIZ ANTONIO 1700 4 ANDAR
5745675,302165	4058010,720	17157546000	17157546	0001	53	CIA DE MINERACAO SERRA DA FAROFA-CEFAR	PRACA NAMI JAFET 235
5745675,302165	4058010,720	17157546000	17157546	0001	53	CIA DE MINERACAO SERRA DA FAROFA-CEFAR	PRACA NAMI JAFET 235
5745675,302165	4058010,720	60599370000	60599370	0001	19	EMPRESA CONTINENTAL DE MINERIOS S/A	PRACA NAMI JAFET 000235
5745675,302165	4058010,720	60727120000	60727120	0001	17	MINERACAO GERAL DO BRASIL LTDA	PRACA NAMI JAFET 235
5735916,721932	4061639,461	76662071000	76662071	0001	11	MINERACAO SAO BRAZ S/A.	RUA EVEZU 119
5736209,816589	4062410,319	00573193000	00573193	0001	70	MINERACAO ITAPETININGA LTDA.	RUA GUAIANAZ DA FONSECA 000028
5742048,94773	4061493,688	57631285000	57631285	0001	86	COMVALER MINERACAO EXP E IMP LTDA	RUA MAJOR QUEDINHO 000111 21 ANDAR
5763986,635175	4066098,467	00288739000	00288739	0001	40	AMERICO BORGES SANTOS FILHO AVICOLA ME	RUA CEMBIRA 001040
5755389,42147	4057712,015	53487153000	53487153	0001	71	AVICOLA CORUJAO LTDA - ME	PRACA EULALIA DE CARVALHO 000004
5759532,581538	4067779,623	74101569000	74101569	0007	75	AGRICOLA JANDELLE LTDA.	RUA JOSE LOPES RODRIGUES 29 000029

Tabela no formato DBASE

4.2.3 Dados Tabulares - exemplos



Representação de estruturas de tabelas no Access

4.2.4 Entrada de Dados via Mesa Digitalizadora

Digitalização = conversão de dados analógicos em digitais

- A mesa digitalizadora é um equipamento cuja superfície está associada a uma malha numerada, a qual pode ser relacionada a um sistemas de coordenadas.

4.2.4 Entrada de Dados via Mesa Digitalizadora

- Formas de digitalização:
  - Passo a passo: digitalização das feições do mapa através da captura intercalada de pontos com o cursor.
  - Contínua: digitalização das feições através da extração contínua do contorno da feição com o cursor.

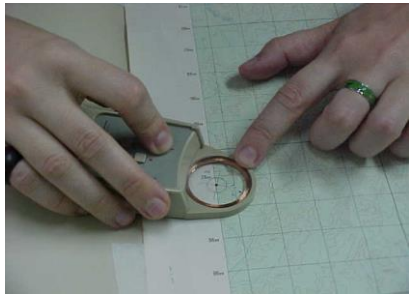
- Etapas da digitalização via Mesa digitalizadora:

1. Posicionar mapa na mesa;
2. Entrada das coordenadas de canto do mapa (calibração);
3. Digitalização das linhas (na forma pontual ou contínua);
4. Edição de erros.

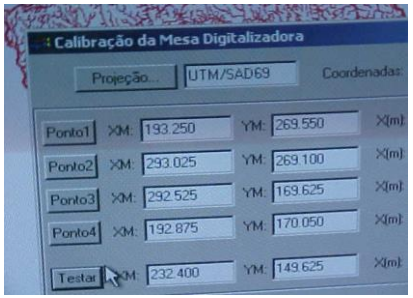
4.2.4. Etapas da Digitalização via Mesa Digitalizadora

Preparo da Mesa

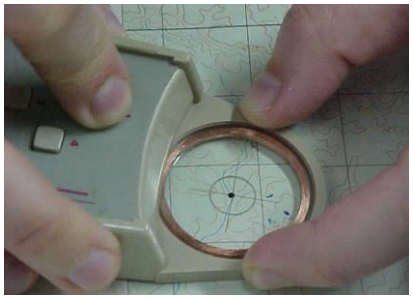
Afixação do mapa à mesa



Calibração da mesa através da coleta de pontos de controle



Análise do resultado alcançado na calibração da mesa



Digitalização das Feições no mapa

4.2.4 Principais problemas na digitalização

- Distorções dos mapas analógicos: dobras e borrões
- Erros inerentes ao mapa: escala e simbologia inadequadas
- Erros do operador
- Resolução final da informação

4.2.5 Georreferenciamento dos dados

- Os dados digitalizados podem ser georreferenciados antes de serem inseridos no SIG.

- O georreferenciamento consiste numa transformação geométrica que relaciona coordenadas da imagem (linha e coluna) com um sistema de coordenadas conhecido.

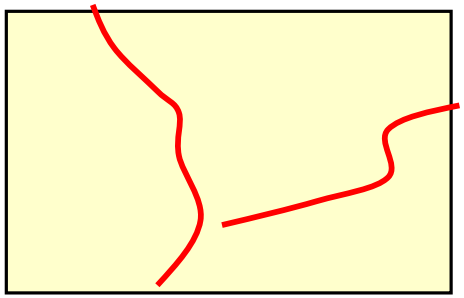
- Essa transformação permite a comparação e a sobreposição com qualquer outro dado espacial, desde que igualmente georreferenciado.

4.2.5 Georreferenciamento dos dados

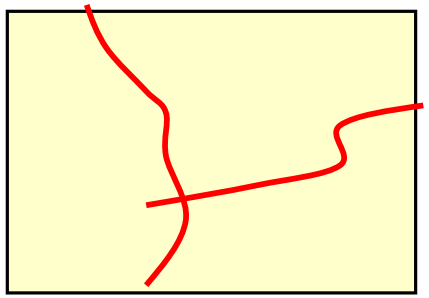
Por que fazer o georreferenciamento?

- Integração de imagens obtidas por sensores diferentes
- Integração de imagens obtidas em tempos diferentes (análise temporal)
- Integração de imagens tomadas em posições diferentes (informação tridimensional)
- Mosaico de imagens

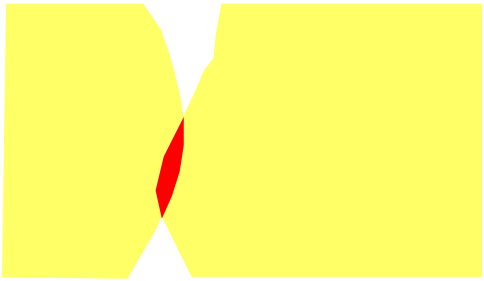
4.2.6. Erros de medição de dados
Exemplos de Erros Comuns



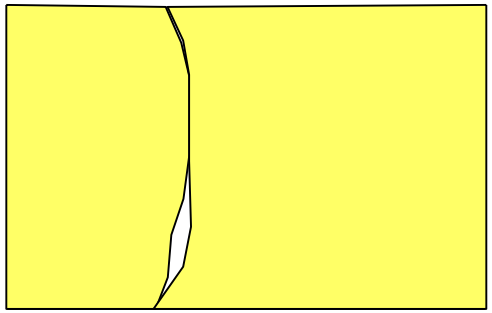
Não Conexão entre linhas (*undershoot*)



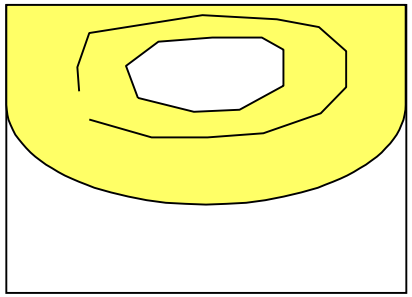
Linha se sobrepõe a outra (*overshoot*)



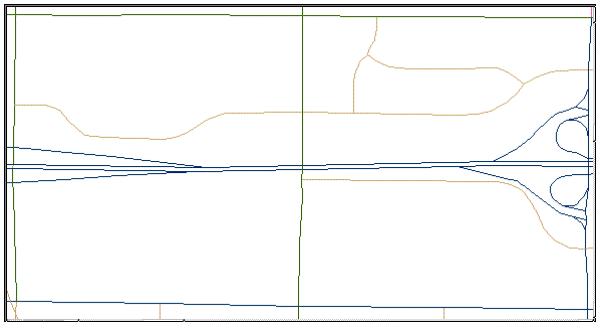
Polígonos sobrepostos (slivers)



Espaço entre polígonos (gaps)



Polígono não fechado



Descontinuidade nas margens

2.4.6 Erros de medição de dados

- Alguns erros podem ser corrigidos automaticamente:
  - Pequenos espaços nas junções de polígonos ou linhas;
  - Linhas que ultrapassam outras linhas (overshoot)
  - Linhas que não se conectam (undershoot)
  - Sobreposição de polígonos

- Esta fase constitui duas importantes etapas:
  - Ajuste de linhas;
  - Validação Topológica

2.4.6 Erros de medição de dados

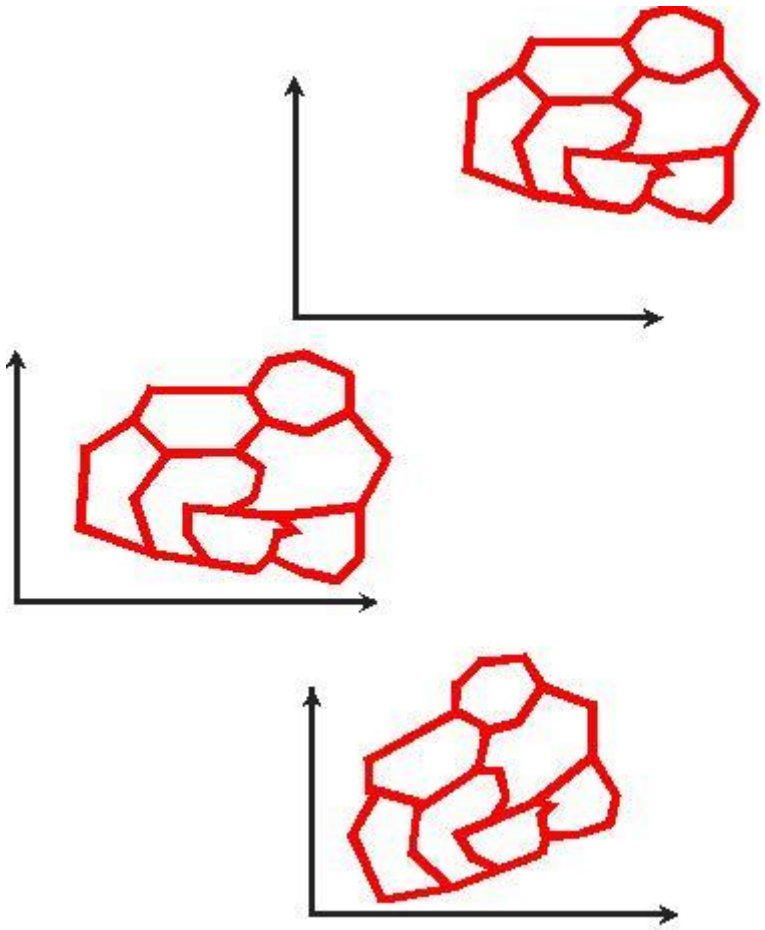
- Erros na entrada de dados podem ocorrer de três fontes principais:
  - nos dados originais
  - durante a inserção
  - durante a transferência e a conversão de dados

- A gerência contínua da qualidade dos dados é necessária neste estágio em que se fará a identificação e a correção dos erros

2.4.6 Erros de medição de dados

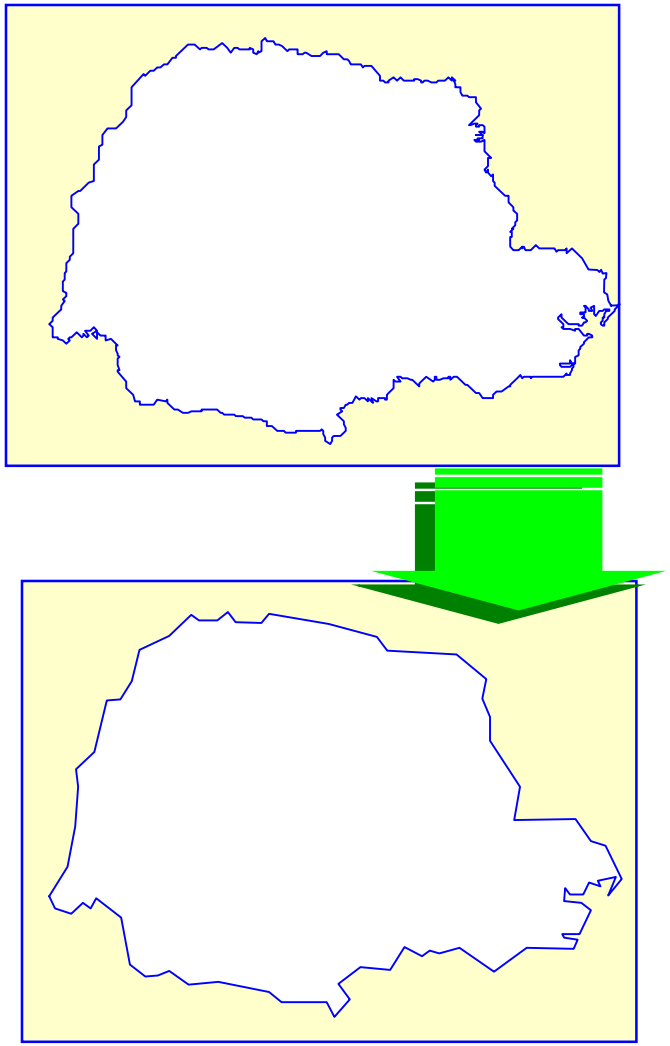
- Detecção e correção dos erros
  - erro de digitação e/ou digitalização
  - codificação errada de uma linha
  - dobras ou manchas do mapa em papel podem ser confundidas com características geográficas reais
- Transformação de coordenadas
  - projeções diferentes da usada num determinado SIG
  - mapas de diferentes escalas
- Generalização
  - mapas de diferentes escalas

Transformação de Coordenadas



- projeções diferentes
- transformação matemática
  - criação de uma origem comum
  - rotação
  - tradução e “scaling”

Generalização



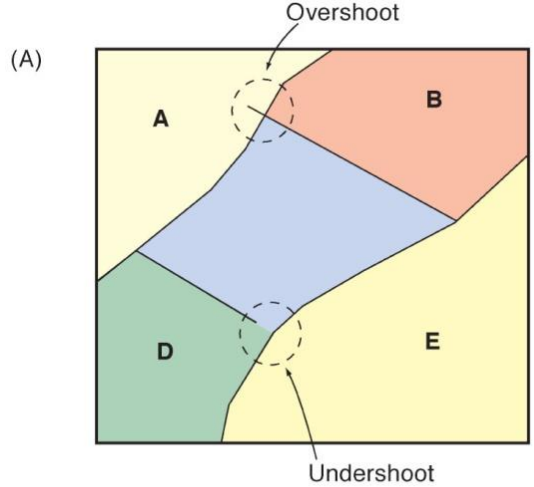
- remover pontos desnecessários
- Exclusão de pontos ao longo de uma linha em um intervalo fixo

Edição de dados

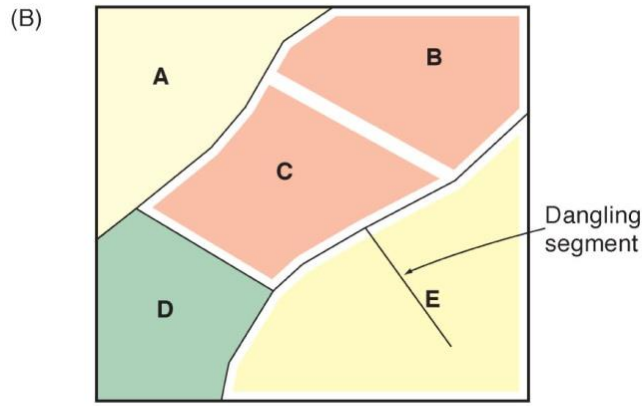
- erros em atributos
  - valores impossíveis
  - valores extremos
  - consistência interna

- erros em dados espaciais
  - feições faltantes
  - feições duplicadas
  - identificador perdido
  - identificador duplicado

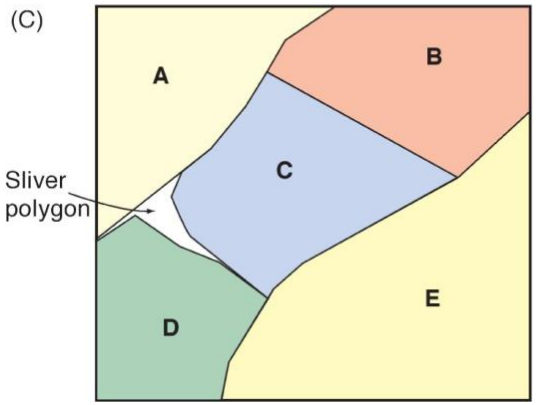
Examples of human errors in digitizing



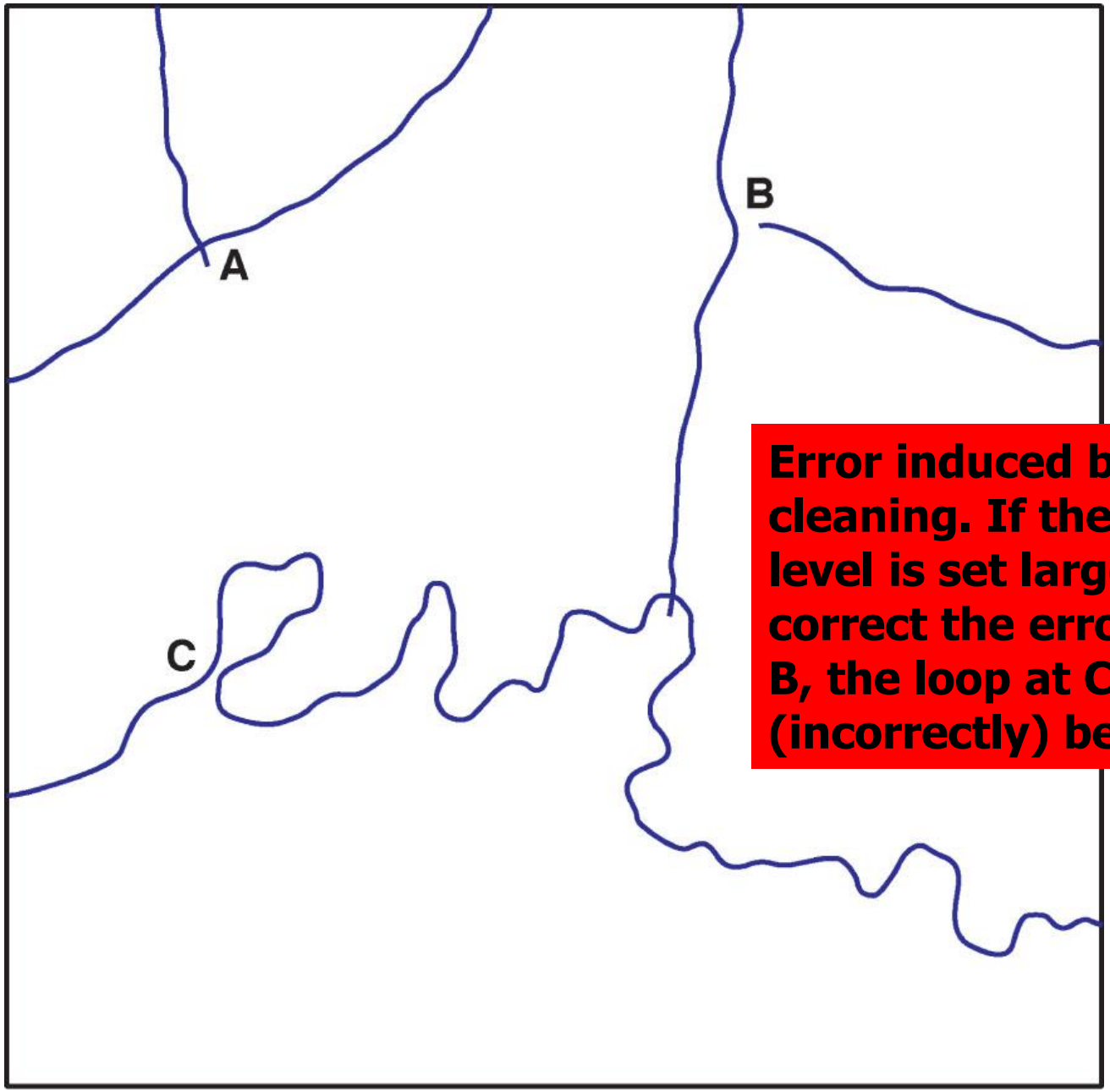
(A) undershoots and overshoots



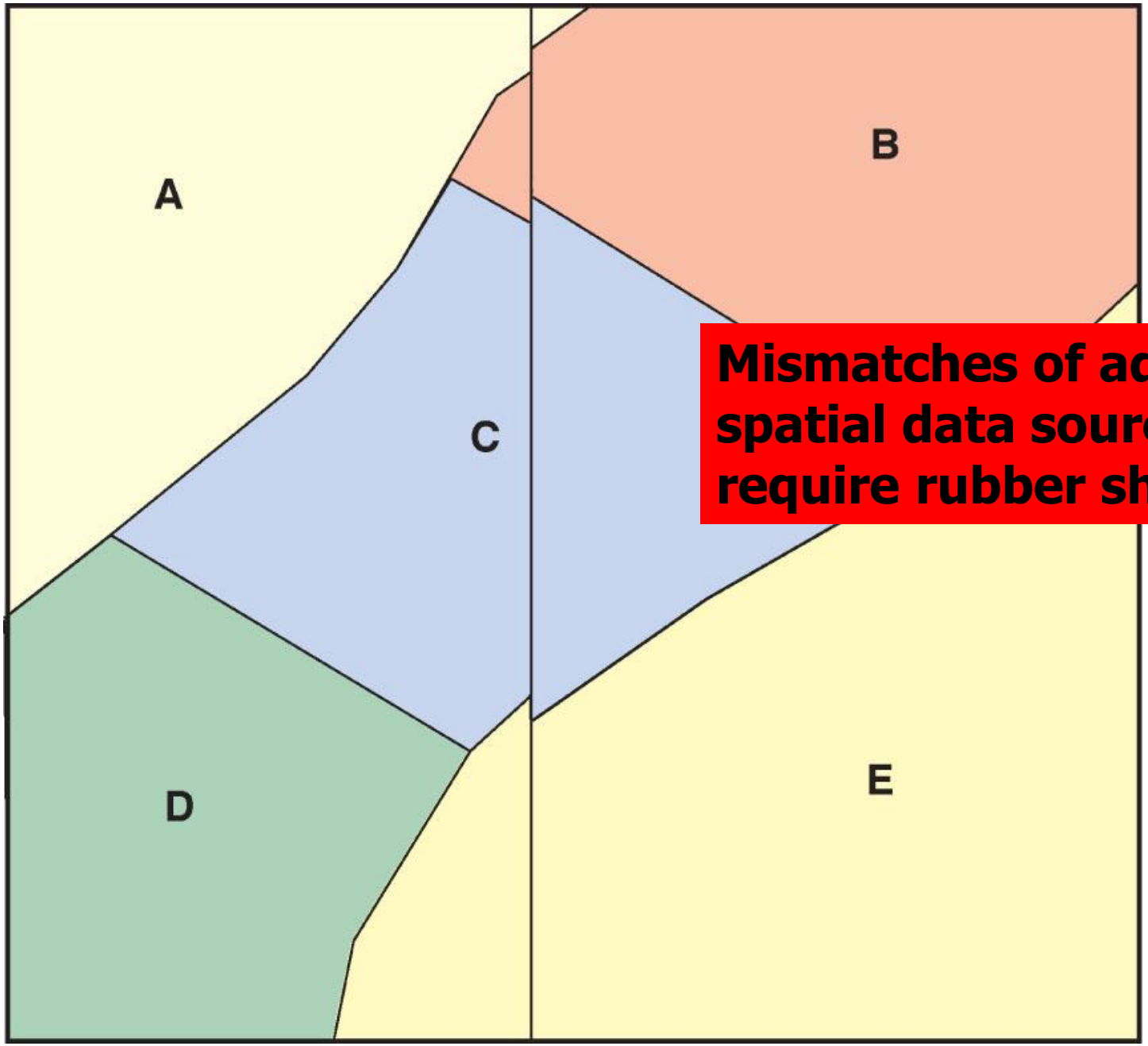
(B) invalid polygons



(C) sliver polygons



Error induced by data cleaning. If the tolerance level is set large enough to correct the errors at A and B, the loop at C will also (incorrectly) be closed.



Mismatches of adjacent spatial data sources that require rubber sheeting

5. Obtenção de dados de fontes externas (transferência de dados)

- Características:
  - Processo de importação (transferência de dados) para um SIG de dados produzidos ou capturados por terceiros
  - Dados disponíveis gratuitamente ou pagos
  - Dados podem ser encontrados em Infraestruturas de Dados Espaciais (IDE):
    - Geobibliotecas
    - Geoportais
  - Possibilidade de busca de dados por metadados

LONGLEY et al. (2013), pg. 244

Considerações finais

- É preciso cuidado nos processos de seleção, entrada e conversão de dados a serem empregados para a formação de uma base em SIG.
- Os erros gerados nestas fases podem prejudicar a qualidade dos dados e fazer com que a confiabilidade das informações obtidas seja afetada.
- O tempo gasto no planejamento destas etapas reflete no resultado a ser alcançado.

Questões para estudo

LONGLEY, PAUL A.; GOODCHILD, MICHAEL, F.; MAGUIRE, DAVID, J.; RHIND, DAVID, W. **Sistemas e Ciência da Informação Geográfica**. 3ª ed. Bookman, Porto Alegre, 2011, 539p

Capítulo 9: Coleta de Dados para SIG [TÉCNICAS]

Questão 1 (modificada): Quais são as vantagens da vetorização automática em lote sobre a vetorização manual em tela? Cite quais tipos de erros podem ser evitados.

Questão 2 (modificada): Que medidas de garantia de qualidade você incluiria em um projeto de coleta de dados desenvolvido para construir um banco de dados de lotes para a avaliação fiscal? Quais tipos de erros devem ser endereçados?