



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

**MODELOS, LINGUAGENS E ARQUITETURAS
PARA
BANCOS DE DADOS GEOGRÁFICOS**

Tese submetida ao Programa de Doutorado em
Computação Aplicada como requisito parcial para o título de Doutor
por
GILBERTO CÂMARA
SAO JOSÉ DOS CAMPOS, SP
DEZEMBRO, 1995

PARECER DA BANCA EXAMINADORA

Este documento foi aprovado como versão final da tese de Doutorado de Gilberto Câmara.

São José dos Campos, SP, 20 de dezembro de 1995.

Presidente da Banca:

Dr. Tatuó Nakanishi, pesquisador, INPE

Orientador:

Dr. Marco Antonio Casanova, pesquisador, IBM Brasil

Convidada externa:

Dra. Cláudia Bauzer Medeiros, professora, UNICAMP

Convidado externo:

Dr. Geovane Cayres Magalhães, professor, UNICAMP

Convidada externa:

Dra. Andrea Silva Hemerly, pesquisadora, IBM Brasil

Examinador:

Dr. Solon Venâncio de Carvalho, pesquisador, INPE

Examinador:

Dr. Diógenes Salas Alves, pesquisador, INPE

“Trazendo de países distantes nossas formas de convívio, nossas instituições, nossas idéias, e timbrando em manter tudo em ambiente muitas vezes desfavorável e hostil, somos ainda hoje uns desterrados em nossa terra. Podemos construir obras excelentes, enriquecer nossa humanidade de aspectos novos e imprevistos, elevar à perfeição o tipo de civilização que representamos; o certo é que todo o fruto de nosso trabalho ou de nossa preguiça parece participar de um sistema de evolução próprio de um outro país e de outra paisagem”.

S.B. de Hollanda, “Raízes do Brasil”.

Para os colegas Ubirajara Freitas e Ricardo Cartaxo,
pela insistência em construir um futuro,
sem os quais este trabalho não teria sentido.

“Au pays parfumé que le soleil caresse,
J’ai connu, sous un dais d’arbres tout empourprés
Et de palmiers d’où pleut sur les yeux la paresse,
Une dame créole aux charmes ignorés.

Son teint est pâle et chaud; la brune enchanteresse
A dans le cou des airs noblement maniérés;
Grande et svelte en marchant comme une chasseresse,
Son sourire est tranquille et ses yeux assurés.

Si vous alliez, Madame, au vrai pays de gloire,
Sur les bords de la Seine ou de la verte Loire,
Belle digne d’orner les antiques manoirs,

Vous feriez, à l’abri des ombreuses retraites,
Germer mille sonnets dans le coeur des poètes,
Que vos grands yeux rendraient plus soumis que vos noirs.”

Baudelaire, “Les fleurs du mal”.

Para Vera Lúcia.

AGRADECIMENTOS

O desenvolvimento desta tese é uma consequência direta da missão da Divisão de Processamento de Imagens do INPE em dotar o Brasil de competência na área de Geoprocessamento. Sem a existência de um grande esforço no desenvolvimento tecnológico, que permita a apropriação social do conhecimento científico, não há sentido em perseguir metas puramente acadêmicas.

Deste modo, agradeço em primeiro lugar aos companheiros da DPI, que garantem com grande denodo a geração de tecnologia na área de sistemas de informação geográfica e de processamento de imagens para a sociedade brasileira.

Esta tese também não teria acontecido sem a cooperação da equipe do Centro Científico Rio da IBM Brasil. A competência aportada ao INPE pela equipe do CC-Rio vem sendo de enorme valia.

Importa aqui destacar a figura de Marco Casanova, orientador deste trabalho e pessoa de enorme competência e de fino trato, mentor intelectual e incentivador deste trabalho.

Agradeço ainda muito a Ubirajara Freitas, que pacientemente revisou trechos importantes desta tese, e que contribuiu com muitas idéias e sugestões. Seu estímulo constante foi fundamental para que este trabalho se completasse.

Contribuição importante também foi a de Diógenes Alves, que realizou um trabalho pioneiro no Brasil sobre o problema de linguagens de consulta e manipulação espacial, e com quem sempre tive discussões esclarecedoras.

Finalmente, agradeço a meu companheiro Ricardo Cartaxo, poço de “insights” e dotado de intuição fora do comum, sempre capaz de fazer o comentário pertinente.

PREFÁCIO

O engajamento num programa de Doutorado foi uma decisão relativamente tardia em minha carreira de engenheiro no INPE. Apenas após completar dez anos de Mestrado, e depois de haver exercido cargos de gerente de projetos como o SGI, MicroMAGICS e SPRING, é que resolvi encarar o desafio de vontade e perseverança envolvido na elaboração de uma tese.

Olhando em retrospecto, considero que o processo de maturação foi extremamente benéfico. Em 1982, quando completei o Mestrado, a área de Geoprocessamento (aonde eu viria a me especializar) estava apenas começando no Brasil. Tenho inclusive um mal-disfarçado orgulho em haver sido o primeiro a publicar trabalhos sobre o assunto em congressos da SBC e sempre lembro que começamos a desenvolver o SGI na mesma época que seus similares no estrangeiro. Hoje, quando desde pequenas prefeituras até grandes concessionárias de serviços públicos estão instalando sistemas de Geoprocessamento, considero que foi fundamental ter-me envolvido no desenvolvimento e uso desta tecnologia, para compreender todo seu contexto, antes de realizar uma contribuição acadêmica de maior vulto.

Outro aspecto relevante foi a escolha do orientador. Sempre considerei que deveria evitar fazer uma tese de Doutorado na mesma área que o Mestrado, para não ficar limitado em meus horizontes profissionais. Assim, após concluir o Mestrado em Processamento de Imagens com o professor Nélson Mascarenhas, excelente orientador e grande figura humana, decidi que o Doutorado deveria ser em outra área, a qual nosso grupo no INPE era reconhecidamente carente: a junção das tecnologias de Bancos de Dados com a de Geoprocessamento. Tive de esperar 10 longos anos até achar o orientador que precisava, mas valeu a pena.

Restava o tema de tese, dilema eterno ao qual sempre nos questionamos, mesmo com o documento quase pronto. Tese abrangente ou específica ? Com a indulgência do orientador, decidi partir para encarar o desafio de “passar a limpo” um conjunto de questões importantes em Geoprocessamento. Que melhor maneira de utilizar a experiência acumulada haveria do que abordar os problemas de formulação conceitual ? Espero contar com a concordância do leitor neste ponto.

Finalmente, lugar-comum tão inevitável quanto merecedor, gostaria de agradecer a minha esposa, Vera Lúcia e a minha filha, Anita. Como este trabalho foi totalmente escrito em horário reservado à família, somente com uma excepcional dose de compreensão por parte delas pude completá-lo. A elas, muito obrigado.

RESUMO DA TESE

O objetivo desta tese é enfocar o problema de conceber e implementar bancos de dados geográficos. O assunto é aqui tratado de forma abrangente, inclusive com discussões sobre os temas de modelagem conceitual, linguagens de consulta e manipulação e arquiteturas distribuídas. Este tratamento objetiva fornecer uma base conceitual sólida para o desenvolvimento de uma nova geração de sistemas de Geoprocessamento pelo INPE.

O trabalho está dividido em três partes. A primeira parte (capítulos 2 e 3) traz uma apresentação sobre a tecnologia de Geoprocessamento e discute a evolução dos sistemas de informação geográfica. A principal conclusão desta parte é a emergência de um novo paradigma na área de SIG, que envolve avanços em sistemas distribuídos, linguagens de consulta e manipulação espacial.

Os aspectos conceituais de *modelagem* e *linguagens* são o objeto da segunda parte da tese. O capítulo 4 apresenta um modelo conceitual orientada-a-objetos para Geoprocessamento. O capítulo 5 contém uma caracterização das operações em SIGs, de forma coerente com o modelo de dados proposto. O Capítulo 6 propõe então uma linguagem para análise geográfica.

A terceira parte da tese analisa arquiteturas para sistemas de gerência de bancos de dados geográficos (SGBDG). O capítulo 7 apresenta uma análise das alternativas gerais de arquiteturas para SGBDG. Com base nos requisitos e na álgebra de operações proposta, o capítulo contém uma proposição de uma arquitetura distribuída. O capítulo 8 discute os problemas envolvidos em implementar a arquitetura proposta e otimizar seu desempenho. Esta parte detalha alternativas para melhoria de desempenho e realiza análises comparativas.

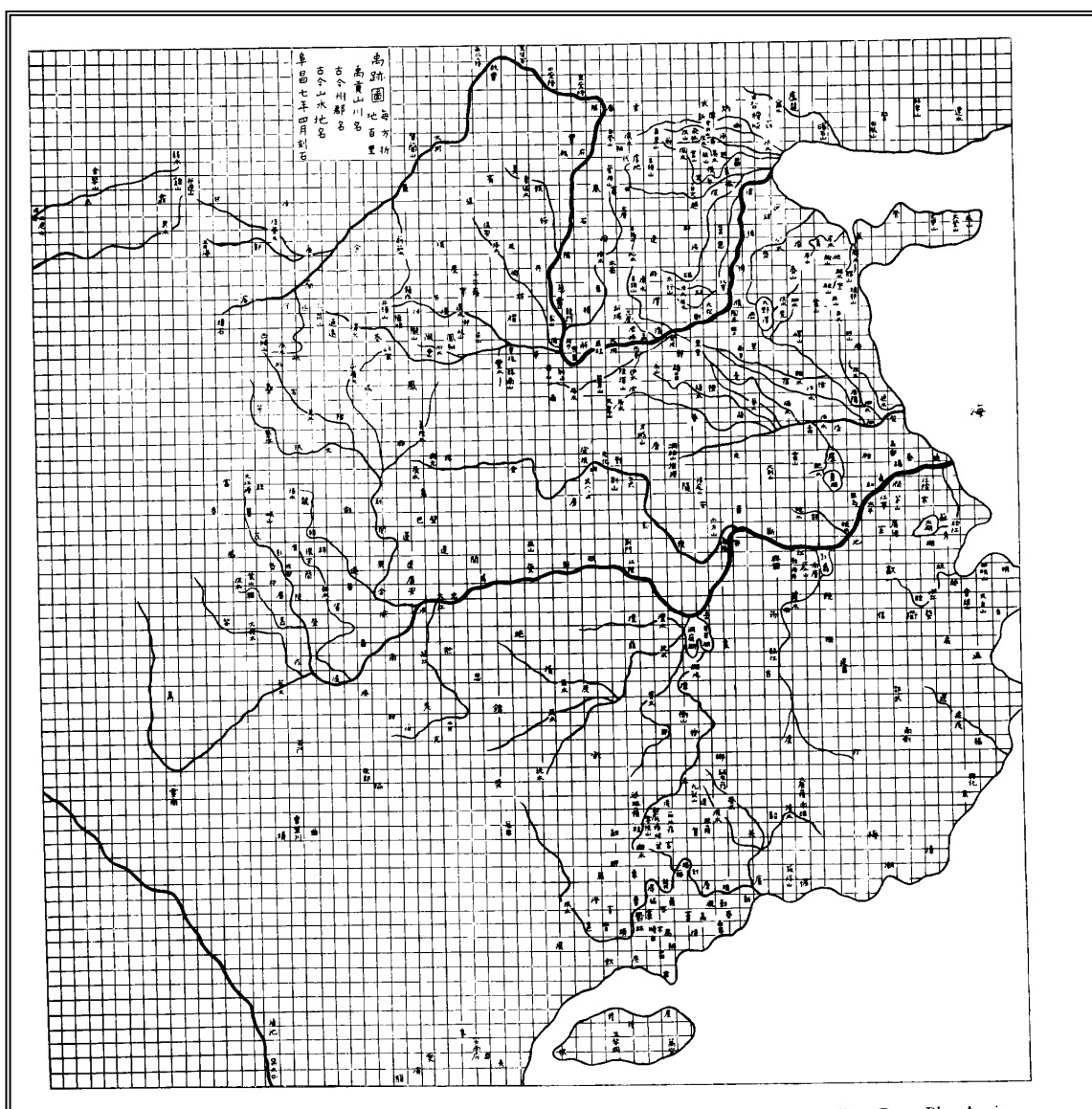
ABSTRACT

The aim of this thesis is to study the issues of design and implementation of geographical data bases. The subject is dealt with in a comprehensive manner, including discussions on the issues of modelling, query and manipulation languages and distributed architectures. Such an approach is meant to give a solid conceptual basis for the development of the next generation of geographical information systems (GIS) at INPE.

The work is divided into three parts. The first part (chapters 2 and 3) includes a general presentation of GIS technology and discusses its evolution. The main conclusion is the emergence of a new paradigm in this arena, based on advances on the areas of distributed systems and query and manipulation languages.

Data modelling and languages are the subject of the second part of this document. Chapter 4 presents an object-oriented data model for geographical data, which serves as a basis for the classification of geographical operations, the subject of chapter 5. Chapter 6 proposes a geographical query and manipulation language, which is based on this classification.

The third part of the document covers architectures for geographical data base management systems (GDBMS). Chapter 7 analyses alternatives for GDBMS architectures. Based on the requirements for the new generation of GISs (chapter 3) and on the language for GIS operations, the chapter proposes a distributed architecture. Chapter 8 discusses the problems involved in implementing the proposed architecture.



O Mapa das Trilhas do Grande Yü, descrito como “o mais remarcado trabalho cartográfico da sua época, talhado na pedra por volta do ano 1100. A linha de costa é firme e a precisão da rede hidrográfica extraordinária. A escala da grade é de 100 li por divisão. O original, que agora está no Museu Pei Lin em Xi'an, tem cerca de 1 metro quadrado.” (Citado por Needham, “Science and Civilization in China”). Reproduzido de Tufte (1983).

1 INTRODUÇÃO

“O futuro não precisa apenas ser imaginado: precisa ser construído. O arquiteto precisa ser capaz de sonhar com coisas que ainda não foram criadas - uma catedral aonde hoje é apenas um terreno empoeirado. Mas o arquiteto também precisa ser capaz de gerar uma planta que mostre como transformar o sonho em realidade.”

*G. Hampel e C.K. Pralahad
(“Competindo pelo Futuro”)*

1.1 MOTIVAÇÃO

A tecnologia de Geoprocessamento irá se converter numa das principais ferramentas de gerenciamento dos recursos naturais e urbanos do Brasil até o final da década. A partir da Constituição de 1988, que consagrou noções como a de crime ecológico, e da RIO '92, os problemas ambientais saíram do âmbito restrito a grupos ecológicos e ONGs para adquirir status de questões nacionais. Pode-se perguntar: estamos diante de uma moda passageira ou estas questões vieram para ficar ?

Em nossa opinião, estes problemas são permanentes e tendem a crescer em relevância nas próximas décadas. No contexto de um novo projeto de desenvolvimento para o Brasil, o tratamento de nosso meio-ambiente, o planejamento agrícola e o zoneamento ecológico-econômico são temas presentes de qualquer agenda política.

Durante a década de 80, o Brasil foi capaz de desenvolver a tecnologia de Geoprocessamento, em níveis de competência em tudo comparáveis às soluções internacionais. Liderados pelo INPE, que construiu o software SGI/SITIM e a tecnologia de hardware necessária, instituições nacionais conseguiram resultados muito significativos na área (veja-se, por exemplo, Assad e Sano, 1993). A disponibilidade de sistemas SGI/SITIM foi essencial para criar competência

nacional em Sensoriamento Remoto e Geoprocessamento no País (Souza et. Al., 1990).

Ao contrário do que possa parecer, o sucesso do SGI/SITIM não se deve apenas à reserva de mercado em informática, mas também ao fato de que seu desenvolvimento ocorreu na forma e momento adequados. Ao adotar a plataforma PC como base de trabalho, a equipe do INPE apostou corretamente em fornecer aos usuários do País uma ferramenta de fácil utilização e que permitiu disseminar a tecnologia de Geoprocessamento.

Nesta década de 90, estamos vivendo um cenário caracterizado por uma crescente inserção do Brasil na economia internacional. Este panorama impõe um grande desafio a todas as equipes de pesquisa e desenvolvimento no País envolvidas na geração de tecnologia: *é preciso manter um contínuo processo de inovação e garantir que os resultados possam chegar aos usuários, sob forma de produtos.*

Neste contexto se insere o presente trabalho. Trata-se de *compreender* bem a evolução da tecnologia, *formalizar* as operações de consulta, manipulação e apresentação, e *projetar* ambientes adequados para a nova geração de sistemas de Geoprocessamento.

1.2 OBJETIVO DO TRABALHO

Esta tese procura mapear os requisitos de evolução da tecnologia de Geoprocessamento e estabelecer uma *visão sistêmica* do problema do projeto de Bancos de Dados Geográficos. O objetivo fundamental deste trabalho é ajudar a projetar uma nova geração de sistemas para Geoprocessamento e assim contribuir para a evolução da sistema SPRING.

Além de objetivos tecnológicos e comerciais, o SPRING é uma plataforma para novas pesquisas em Geoprocessamento, Processamento de Imagens e Bancos de Dados Geográficos. A base do sistema é uma biblioteca de classes em C++, que encapsula métodos de acesso e de processamento para dados geográficos e facilita

extensões, como o uso de sistemas de gerência de bancos de dados orientados-a-objetos. Para maiores detalhes sobre o SPRING, veja-se Câmara et al. (1992) e Câmara et al (1993).

Além do desenvolvimento do SPRING como ferramenta de trabalho para projetos em Geoprocessamento, o INPE estará implantando, até o final da década, um Centro de Dados de Sensoriamento Remoto. Este centro deverá tornar disponível à comunidade um amplo conjunto de dados geográficos e ambientais, resultado de projetos desenvolvidos pelo INPE e instituições conveniadas. Apoiar a evolução da tecnologia SPRING para atender a este desafio é também um dos objetivos desta tese.

1.3 OS DESAFIOS CIENTÍFICOS

O desenvolvimento da tecnologia de Geoprocessamento vem acontecendo de forma paralela e em grande medida independente da evolução recente da Informática. Como resultado, há uma defasagem entre os avanços do setor de Computação e seu uso da tecnologia de Geoprocessamento e a maior parte das soluções comerciais no setor possui limitações conceituais.

Para acentuar o contraste, notemos que não há até o momento na área um equivalente ao trabalhos de Codd sobre álgebra relacional (Codd, 1970) e que pudessem resultar num padrão como a linguagem SQL. Para superar esta e outras deficiências, o programa de pesquisas da comunidade de Geoprocessamento procura resolver questões como:

- Como projetar uma álgebra de operações geográficas que seja suficientemente rica e que possa dar origem a uma linguagem de consulta e manipulação padrão para bancos de dados geográficos ?
- Como integrar dados de sensoriamento remoto num sistema de informações geográficas ?
- Como construir ambientes com capacidade de gerenciar *grandes bancos de dados geográficos* ?

- Como integrar processos de *modelagem* que permitam ter uma visão dinâmica dos processos que afetam o ambiente ?
- Como tornar disponíveis técnicas que permitam os dados produzidos por um SIG possam operar de forma coordenada com outros produtos ?

1.4 ORGANIZAÇÃO DA TESE

Inicialmente é feita, no capítulo 2, uma apresentação sobre a tecnologia de Geoprocessamento. Esta discussão tem caráter meramente informativo e objetiva fornecer um pano de fundo para o trabalho.

A seguir, o capítulo 3 apresenta uma visão sobre as perspectivas de evolução tecnológica, é apresentado um modelo de dados para Geoprocessamento. A tecnologia é dividida em sistemas de primeira geração (ditos “CAD cartográfico”), segunda (“bancos de dados geográficos”) e terceira (“bibliotecas digitais”). A metodologia apresentada, apesar de consensual entre os especialistas¹, é ainda inédita na literatura.

No capítulo 4, apresentamos um modelo de dados orientado-a-objetos para Geoprocessamento, cujas características principais são: separar o universo conceitual do universo de implementação e suportar as noções de *campos* e *objetos*.

A partir deste modelo, são formalizadas no capítulo 5 as operações de consulta, manipulação e apresentação de dados geográficos. O capítulo 6 contém uma proposta de uma linguagem de consulta e manipulação de dados geográficos, chamada LEGAL.

No Capítulo 7, fazemos uma análise das diferentes arquiteturas para gerência de dados geográficos e o elaboramos uma proposta de uma arquitetura distribuída para sistemas de Geoprocessamento. Completa o trabalho, no Capítulo 8, uma

¹Em conversa com Max Egenhofer (talvez o maior teórico atual em Geoprocessamento), ele concordou com a visão expressa no capítulo 3.

discussão sobre a implementação da linguagem LEGAL na arquitetura distribuída proposta.

1.5 CONTRIBUIÇÃO DA TESE

Esta tese nasceu de um trabalho coletivo, feito pela Divisão de Processamento de Imagens do INPE em convênio com o Centro Científico Rio da IBM e baseia-se nos mais recentes trabalhos da literatura.

Sem embargo, há vários pontos no trabalho que representam (na expectativa do autor e seu orientador) contribuição inédita significativa:

- O modelo abrangente orientado-a-objetos para sistemas de informação geográfica (exposto no capítulo 4).
- A visão integradora das operações de Geoprocessamento, proposta no capítulo 5.
- A definição de uma linguagem para consulta e manipulação de dados geográficos (capítulo 6).
- A arquitetura distribuída para bancos de dados geográficos (capítulo 7) e a discussão sobre sua implementação (capítulo 8).

Ademais, esta tese tem o objetivo adicional de apresentar ao leitor uma *visão sistêmica* de problemas atuais em Geoprocessamento. O objetivo foi abordar uma gama relativamente ampla de problemas (desde o modelo até a arquitetura de sistemas, passando pelas operações) e consolidar o conhecimento sobre a nova geração de SIGs.

2 CARACTERÍSTICAS DE SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

“É preciso saber pensar o espaço para saber nele se organizar, para saber ali combater”.

Yves Lacoste

2.1 DESCRIÇÃO GERAL

O termo *sistemas de informação geográfica* (SIG) é aplicado para sistemas que realizam o tratamento computacional de dados geográficos. Devido a sua ampla gama de aplicações, que inclui temas como agricultura, floresta, cartografia, cadastro urbano e redes de concessionárias (água, energia e telefonia), há pelo menos três grandes maneiras de utilizar um SIG:

- como ferramenta para produção de mapas;
- como suporte para análise espacial de fenômenos;
- como um banco de dados geográficos, com funções de armazenamento e recuperação de informação espacial.

Estas três visões do SIG são antes convergentes que conflitantes e refletem a importância relativa do tratamento da informação geográfica dentro de uma instituição. Para esclarecer ainda mais o assunto, apresentamos a seguir algumas definições de SIG:

“Um conjunto manual ou computacional de procedimentos utilizados para armazenar e manipular dados georeferenciados” (Aronoff, 1989).

“Conjunto poderoso de ferramentas para coletar, armazenar, recuperar, transformar e visualizar dados sobre o mundo real” (Burrough, 1986).

“Um sistema de suporte à decisão que integra dados referenciados espacialmente num ambiente de respostas a problemas” (Cowen, 1988).

“Um banco de dados indexados espacialmente, sobre o qual opera um conjunto de procedimentos para responder a consultas sobre entidades espaciais” (Smith et al., 1987)

Estas definições de SIG refletem, cada uma à sua maneira, a multiplicidade de usos e visões possíveis desta tecnologia e apontam para uma perspectiva interdisciplinar de sua utilização. A partir destes conceitos, é possível indicar as principais características de SIGs:

- Integrar, numa única base de dados, informações espaciais provenientes de dados cartográficos, dados de censo e cadastro urbano e rural, imagens de satélite, redes e modelos numéricos de terreno.
- Oferecer mecanismos para combinar as várias informações, através de algoritmos de manipulação e análise e para consultar, recuperar, visualizar e plotar o conteúdo da base de dados geocodificados.

2.2 CONCEITOS DE ESPAÇO E RELAÇÕES ESPACIAIS

O que há de especial com dados espaciais? A partir desta pergunta, podemos procurar desvendar as particularidades dos dados geográficos.

O aspecto mais fundamental dos dados tratados em um SIG é a natureza dual da informação: um dado geográfico possui uma *localização geográfica* (expressa como coordenadas em um mapa) e *atributos descritivos* (que podem ser representados num banco de dados convencional). Outro aspecto muito importante é que os dados geográficos não existem sozinhos no espaço: tão importante quanto localizá-los é descobrir e representar as relações entre os diversos dados. Alguns exemplos dos processos de análise espacial típicos de um SIG estão apresentados na tabela abaixo (adaptada de Maguire, 1991).

TABELA 2.1
EXEMPLOS DE ANÁLISE ESPACIAL

Análise	Pergunta Geral	Exemplo
Condição	“O que está...”	“Qual a população desta cidade?”
Localização	“Onde está...?”	“Quais as áreas com declividade acima

		de 20 ?“
Tendência	“O que mudou...?”	“Esta terra era produtiva há 5 anos atrás ?“
Roteamento	“Por onde ir..?”	“Qual o melhor caminho para o metrô ?”
Padrões	“Qual o padrão....?”	“Qual a distribuição da dengue em Fortaleza?”
Modelos	“O que acontece se...?”	“Qual o impacto no clima se desmatarmos a Amazônia ?”

Tomemos um exemplo concreto para explicitar os conceitos acima sobre análise espacial (neste caso realizada manualmente). Em 1854, Londres estava sofrendo uma grave epidemia de cólera, doença sobre a qual na época não se conhecia a forma de contaminação. Numa situação aonde já haviam ocorrido mais de 500 mortes, o doutor John Snow teve um “estalo”: colocar no mapa da cidade a localização dos doentes de cólera e dos poços de água (naquele tempo, a fonte principal de água dos habitantes da cidade). O mapa obtido está mostrado na figura 2.1.

Com a espacialização dos dados, o doutor Snow percebeu que a maioria dos casos estava concentrada em torno do poço da “Broad Street” e ordenou a sua lacração, o que contribuiu em muito para debelar a epidemia. Este caso forneceu evidência empírica para a hipótese (depois comprovada) de que o cólera é transmitido por ingestão de água contaminada. Este é uma situação típica aonde a relação espacial entre os dados muito dificilmente seria inferida pela simples listagem dos casos de cólera e dos poços.

O mapa do doutor Snow passou para a História como um dos primeiros exemplos que ilustra bem o poder explicativo da análise espacial.

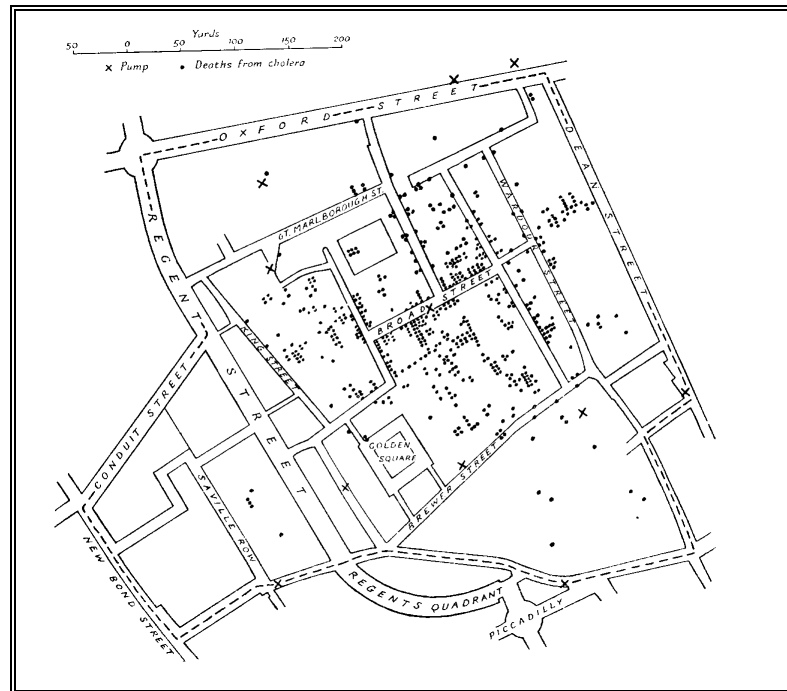


Figura 2.1 - Mapa de Londres com casos de cólera (pontos) e poços de água (cruzes) (adaptado de E. Tufte, 1983).

2.3 ESTRUTURA GERAL DE UM SIG

Numa visão abrangente, pode-se indicar que um SIG tem os seguintes componentes :

- Interface com usuário;
- Entrada e integração de dados;
- Funções de processamento gráfico e de imagens;
- Visualização e plotagem;
- Armazenamento e recuperação de dados (organizados sob a forma de um banco de dados geográficos).

Estes componentes se relacionam de forma hierárquica. No nível mais próximo ao usuário, a *interface homem-máquina* define como o sistema é operado e controlado. No nível intermediário, um SIG deve ter mecanismos de processamento de dados espaciais (*entrada, edição, análise, visualização e saída*). No nível mais interno do sistema, um *sistema de gerência de bancos de dados*

geográficos oferece armazenamento e recuperação dos dados espaciais e seus atributos.

De uma forma geral, as funções de processamento de um SIG operam sobre dados em uma *área de trabalho* em memória principal. A ligação entre os dados geográficos e as funções de processamento do SIG é feita por mecanismos de seleção e consulta que definem restrições sobre o conjunto de dados. Exemplos ilustrativos de modos de seleção de dados são:

- "Recupere os dados relativos à carta de Guajará-Mirim " (*restrição por definição de região de interesse*);
- "Recupere as cidades do Estado de São Paulo com população entre 100.000 e 500.000 habitantes" (*consulta por atributos não-espaciais*).
- "Mostre os postos de saúde num raio de 5 km do hospital municipal de S.J.Campos" (*consulta com restrições espaciais*).

A figura 2.2 indica o relacionamento dos principais componentes. Cada sistema, em função de seus objetivos e necessidades, implementa estes componentes de forma distinta, mas todos os subsistemas citados estão presentes num SIG.

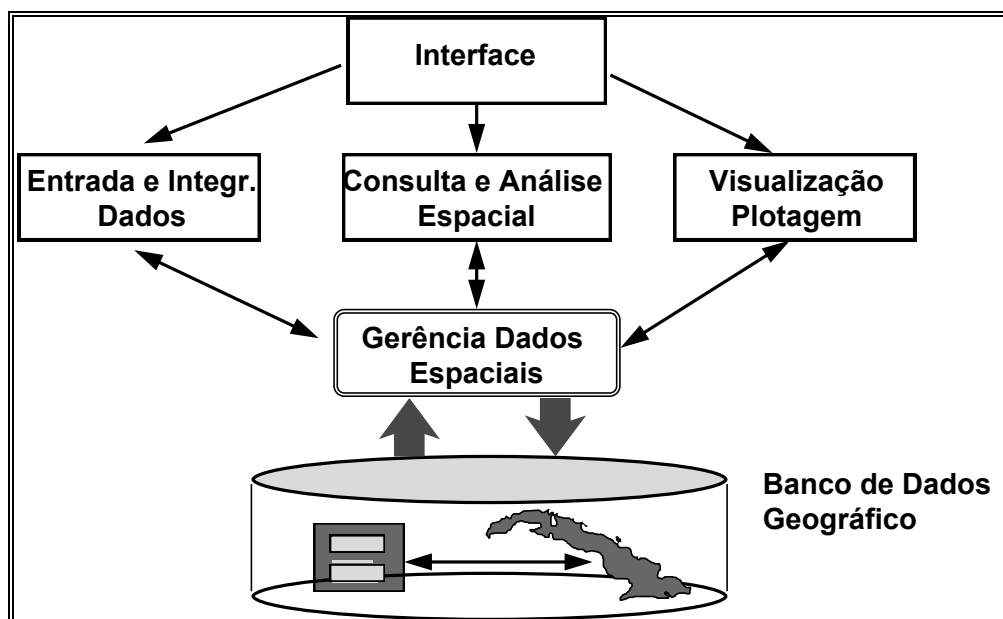


Figura 2.2 - Arquitetura de Sistemas de Informação Geográfica.

2.4 TIPOS DE DADOS EM GEOPROCESSAMENTO

Um sistema de Geoprocessamento armazena a geometria e os atributos dos dados, que estão *georeferenciados*, isto é, localizados na superfície terrestre numa projeção cartográfica. Os dados tratados em geoprocessamento tem como principal característica a diversidade de fontes geradoras e de formatos apresentados.

À diferença de sistemas de Projeto Auxiliado por Computador (CAD), uma característica básica e geral num SIG é sua capacidade de tratar as relações espaciais entre os objetos geográficos. Denota-se por *topologia* a estrutura de relacionamentos espaciais (vizinhança, proximidade, pertinência) que podem se estabelecer entre objetos geográficos.

O requisito de armazenar a geometria dos objetos geográficos e de seus atributos representa uma *dualidade* básica para SIGs. Para cada objeto geográfico, o SIG necessita armazenar seus atributos e as várias representações gráficas associadas.

MAPAS TEMÁTICOS

Mapas temáticos contêm regiões geográficas definidas por um ou mais polígonos. Exemplos são o uso do solo e a aptidão agrícola de uma região. Estes dados, obtidos a partir de levantamento de campo, são inseridos no sistema por digitalização ou, de forma mais automatizada, a partir de classificação de imagens.

Para permitir uma representação e análise mais acurada do espaço geográfico, a maior parte dos sistemas armazena estes tipos de mapas na forma de *arcos* (limites entre regiões), incluindo os *nós* (pontos de intersecções entre arcos) para montar uma *representação topológica*. A topologia construída é do tipo *arco-nó-região*: arcos se conectam entre si através de nós (pontos inicial e final) e arcos que circundam uma área definem um polígono (região).

Um mapa temático também pode ser armazenado no formato matricial ("raster"). Neste caso, a área correspondente ao mapa é dividida em células de

tamanho fixo. Cada célula terá um valor correspondente ao tema mais frequente naquela localização espacial.

A escolha entre a representação matricial e a vetorial para um mapa temático depende do objetivo em vista. Para a produção de cartas e em operações onde se requer maior precisão, a representação vetorial é mais adequada. As operações de álgebra de mapas são mais facilmente realizadas no formato matricial. No entanto, para um mesmo grau de precisão, o espaço de armazenamento requerido por uma representação matricial é substancialmente maior. Isto é ilustrado na figura 2.3.

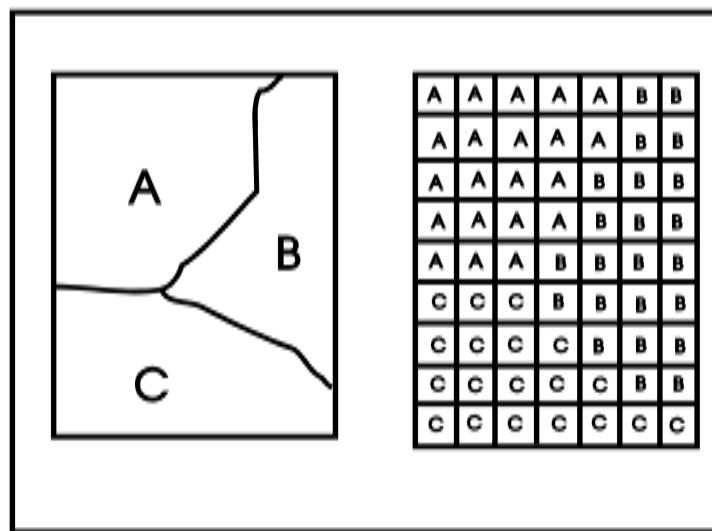


Figura 2.3 - Representação vetorial e matricial de um mapa temático.

A tabela 2.2 abaixo apresenta uma comparação entre as vantagens e desvantagens de armazenamento matricial e vetorial para mapas temáticos. Esta comparação leva em conta os vários aspectos: relacionamentos espaciais, análise, armazenamento. Nesta tabela, o formato mais vantajoso para cada caso é apresentado em destaque.

TABELA 2.2
COMPARAÇÃO ENTRE FORMATOS PARA MAPAS TEMÁTICOS

Aspecto	Formato Vetorial	Formato Varredura
Relações espaciais entre objetos	Relacionamentos topológicos entre objetos disponíveis	Relacionamentos espaciais devem ser inferidos
Ligação com banco de dados	Facilita associar atributos a elementos gráficos	Associa atributos apenas a classes do mapa
Análise, Simulação e Modelagem	Representação indireta de fenômenos contínuos Álgebra de mapas é limitada	Representa melhor fenômenos com variação contínua no espaço Simulação e modelagem mais fáceis
Escalas de trabalho	Adequado tanto a grandes quanto a pequenas escalas	Mais adequado para pequenas escalas (1:25.000 e menores)
Algoritmos	Problemas com erros geométricos	Processamento mais rápido e eficiente.
Armazenamento	Por coordenadas (mais eficiente)	Por matrizes

MAPAS CADASTRAIS

Um mapa cadastral distingue-se de um mapa temático, pois cada um de seus elementos é um *objeto geográfico*, que possui atributos e pode estar associado a várias representações gráficas. Por exemplo, os lotes de uma cidade são elementos do espaço geográfico que possuem atributos (dono, localização, valor venal, IPTU devido, etc.) e que podem ter representações gráficas diferentes em mapas de escalas distintas.

A parte gráfica dos mapas cadastrais é armazenada em forma de coordenadas vetoriais, com a topologia associada. Não é usual representar estes dados na forma matricial.

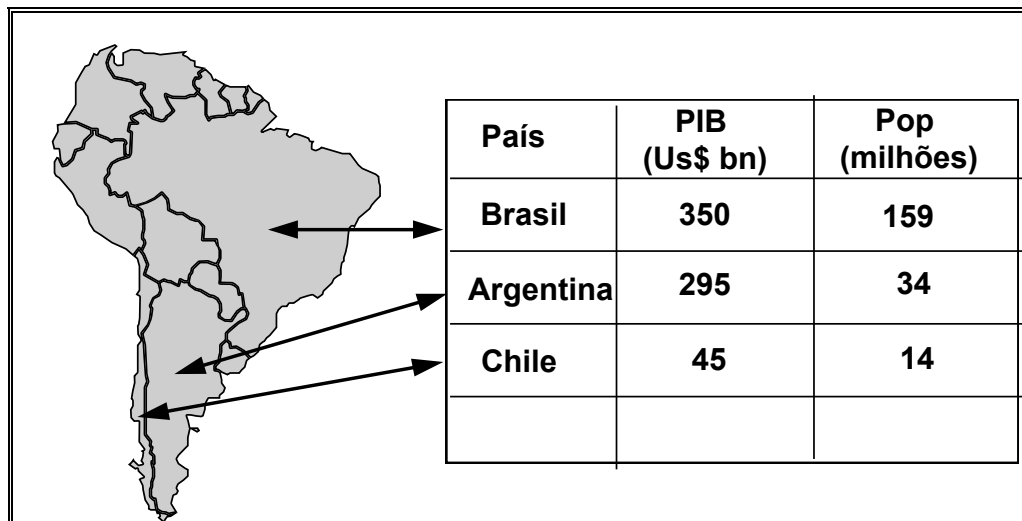


Figura 2.4 - Exemplo de mapa cadastral (países da América do Sul).

REDES

Em Geoprocessamento, o conceito de "rede" denota as informações associadas a:

- Serviços de utilidade pública, como água, luz e telefone;
- Redes de drenagem (bacias hidrográficas);
- Rodovias.

No caso de redes, cada objeto geográfico (e.g: cabo telefônico, transformador de rede elétrica, cano de água) possui uma *localização geográfica exata* e está sempre associado a atributos descritivos, presentes no banco de dados.

As informações gráficas de redes são armazenadas em coordenadas vetoriais, com *topologia arco-nó*: os atributos de arcos incluem o sentido de fluxo e os atributos dos nós sua impedância (custo de percorrimento). A topologia de redes constitui um *grafo*, que armazena informações sobre recursos que fluem entre localizações geográficas distintas, como ilustra a figura 2.5.

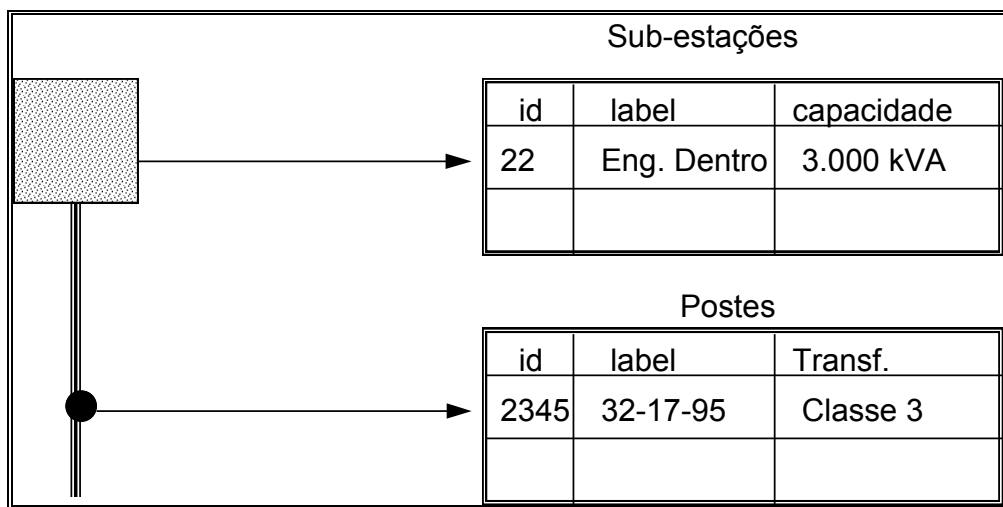


Figura 2.5 - Elementos de Rede.

Como observa Goodchild (1992b), uma rede é um sistema de endereçamento 1-D embutido no espaço 2-D. Para citar um exemplo, tome-se uma rede elétrica, que tem, entre outros, os componentes: postes, transformadores, sub-estações, linhas de transmissão e chaves. As linhas de transmissão serão representadas topologicamente como os arcos de um grafo orientado, estando as demais informações concentradas em seus nós. Note-se que os algoritmos de cálculo de propriedades da rede podem, em sua grande maioria, ser resolvidos apenas com a topologia da rede e de seus atributos.

As redes formam um capítulo à parte na tipologia de SIGs, pois - à diferença dos outros tipos de dados - são o resultado direto da intervenção humana sobre o meio-ambiente. Cada aplicação de rede tem características próprias e com alta

dependência cultural (p.ex., a largura das auto-estradas nos EUA é distinta das usadas em São Paulo).

No caso de aplicações em redes, a ligação com *banco de dados* é fundamental. Como os dados espaciais tem formatos relativamente simples, a maior parte do trabalho consiste em realizar consultas ao banco de dados e apresentar os resultados de forma adequada. A área de redes é ainda um grande motivador para inovações em SIG, merecendo destaque:

- integração de dados
- segmentação dinâmica
- linguagem de visualização
- capacidade de adaptação

A *integração de dados* é necessária para aplicações como redes, onde deseja-se gerar uma base cartográfica contínua a partir de informações dispersas em vários mapas. Usualmente, as redes (elétrica, de telefonia e de água e esgoto) estão interligadas em toda a malha urbana. Poucos sistemas conseguem armazená-las de forma contínua, dando origem a particionamentos que não refletem a realidade e que dificultam a realização de análises e simulações.

Outro aspecto necessário para aplicações de redes é a capacidade de definir diferentes cortes lógicos de uma rede sem ter de duplicar (ou repetir) a estrutura topológica da rede. Por exemplo, ao se asfaltar parte de uma estrada de terra, será preciso atualizar esta informação, sem ter de redigitalizar todas as coordenadas de localização da estrada. Esta capacidade, usualmente denotada por *segmentação dinâmica*, permite separar os diferentes níveis de informação relativos a uma mesma rede.

Como as características dos elementos da rede são armazenadas como atributos em bancos de dados, é necessário dispor de meios para visualizar esta informação. Para tanto, os SIGs devem dispor de *linguagem de apresentação*, que permita controlar a simbologia associada aos componentes da rede, que varia conforme a escala de plotagem.

O pacote mínimo disponível nos sistemas comerciais consiste tipicamente de cálculo de caminho ótimo e crítico. Este pacote básico é insuficiente para a realização da maioria das aplicações, pois cada usuário tem necessidades completamente distintas. No caso de um sistema telefônico, uma questão pode ser: "quais são todos os telefones servidos por uma dada caixa terminal?". Já para uma rede de água, pode-se perguntar: "Se injetarmos uma dada percentagem de cloro na caixa d'água de um bairro, qual a concentração final nas casas?"

Deste modo, um sistema de modelagem de redes só terá utilidade para o cliente depois de devidamente adaptado para as suas necessidades, o que pode levar vários anos. Isto impõe uma característica básica para esta aplicação: *os sistemas devem ser versáteis e maleáveis.*

IMAGENS

Obtidas por satélites, fotografias aéreas ou "scanners" aerotransportados, as imagens representam formas de captura indireta de informação espacial. Armazenadas como matrizes, cada elemento de imagem (denominado "pixel") tem um valor proporcional à **reflectância** do **solo** para a área imageada.

Pela natureza do processo de aquisição de imagens, os objetos geográficos estão contidos na imagem e para individualizá-los, é necessário recorrer a técnicas de foto-interpretação e de classificação automática.

Características importantes de imagens de satélite são: o número de bandas do espectro eletromagnético imageadas (*resolução espectral*), a área da superfície terrestre observada instantaneamente por cada sensor (*resolução espacial*) e o intervalo entre duas passagens do satélite pelo mesmo ponto (*resolução temporal*). A tabela 2.3 apresenta as características gerais dos principais satélites (com seus sensores) disponíveis no Brasil.

TABELA 2.3
CARACTERÍSTICAS DE SATÉLITES RECEBIDOS NO BRASIL

<i>Satélite (família)</i>	<i>Instrumento</i>	<i>Num. bandas</i>	<i>Resolução espacial</i>	<i>Resolução temporal</i>
LANDSAT	MSS	4	80 m	18 dias
	TM	7(1)	30 m	18 dias
SPOT	XS	3	20 m	27 dias ⁽²⁾
	PAN	1	10 m	27 dias
TIROS/NOA A	AVHRR	5	1100 m	6 horas
METEOSAT	MSS	4	8000 m	30 minutos
ERS(3)	SAR banda-C	1	25 m	25 dias

Observações:

1. O sensor TM possui uma banda no infravermelho termal, que tem resolução de 120 metros.
2. O satélite SPOT tem espelho móvel, com visada lateral, o que permite uma capacidade de reduzir o período entre duas tomadas sucessivas de um mesmo ponto da Terra.
3. O satélite ERS-1 é o primeiro a possuir um radar. No entanto, trata-se de satélite experimental, com cobertura limitada.

MODELOS NUMÉRICOS DE TERRENO

O termo *modelo numérico de terreno* (ou MNT) é utilizado para denotar a representação uma grandeza que varia continuamente no espaço. Comumente associados à altimetria, também podem ser utilizados para modelar unidades geológicas, como teor de minerais ou propriedades do solo ou subsolo (como aeromagnetismo).

Entre os usos de modelos numéricos de terreno, pode-se citar (Burrough, 1986):

- (a) Armazenamento de dados de altimetria para gerar mapas topográficos;
- (b) Análises de corte-aterro para projeto de estradas e barragens;
- (c) Cômputo de mapas de declividade e exposição para apoio a análises de geomorfologia e erodibilidade;
- (d) Análise de variáveis geofísicas e geoquímicas;
- (e) Apresentação tridimensional (em combinação com outras variáveis).

O processo de aquisição de uma grandeza com variação espacial produz usualmente um conjunto de amostras pontuais. A partir destas amostras, pode-se construir dois tipos de representação:

- (a) *grades regulares*: matriz de elementos com espaçamento fixo, onde é associado o valor estimado da grandeza na posição geográfica de cada ponto da grade. As grades regulares são obtidas por interpolação das amostras ou, alternativamente, geradas por restituidores com saída digital.
- (b) *malhas triangulares*: a grade é formada por conexão entre amostras, utilizando a triangulação de Delaunay (sujeita a restrições). A grade triangular é uma estrutura topológica vetorial do tipo arco-nó, que forma um conjunto de recortes irregulares no espaço.

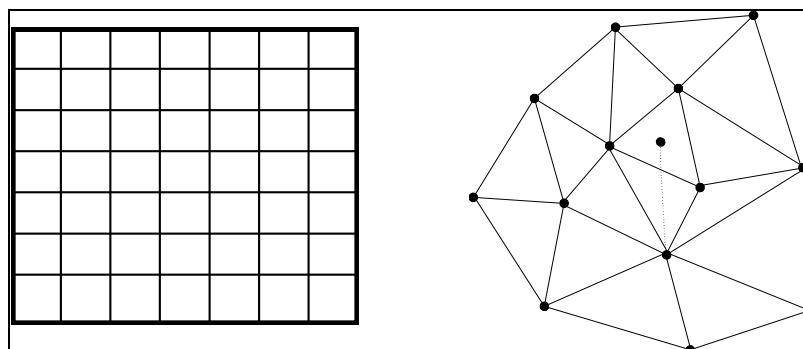


Figura 2.6 - Grades regulares e triangulares.

Os procedimentos de interpolação para geração de grades regulares a partir de amostras variam de acordo com a grandeza medida. No caso de altimetria, é comum o uso de funções de ponderação por inverso do quadrado da distância. Já para variáveis geofísicas, procedimentos de filtragem bidimensional ou de geoestatística (como a *krigeagem*) são utilizados.

As grades triangulares são normalmente melhores para representar a variação do terreno, pois capturam a complexidade do relevo sem a necessidade de grande quantidade de dados redundantes. As grades regulares tem grande redundância em terrenos uniformes e dificuldade de adaptação a relevos de natureza distinta no mesmo mapa, por causa da grade de amostragem fixa.

Para o caso de variáveis geofísicas e para operações como visualização 3D, as grades regulares são preferíveis, principalmente pela maior facilidade de manuseio computacional.

A tabela 2.4 resume as principais vantagens e desvantagens de grades regulares e triangulares.

TABELA 2.4
COMPARAÇÃO ENTRE GRADES REGULARES E TRIANGULARES PARA
REPRESENTAR MODELOS NUMÉRICOS DE TERRENO

	Grade triangular	Grade regular
Vantagens	1. Melhor representação de relevo complexo 2. Incorporação de restrições como linhas de crista	1. Facilita manuseio e conversão 2. Adequada para geofísica e visualização 3D
Problemas	1. Complexidade de manuseio 2. Inadequada para visualização 3D	1. Representação relevo complexo 2. Cálculo de declividade

Os modelos numéricos de terreno também podem ser convertidos para mapas temáticos e para imagens. Em ambos os casos, a grandeza numérica é quantizada, seja para um número pequeno de valores (caso de mapas temáticos) seja para a variação associada a imagens (valores discretos).

2.5 CLASSES DE PROJETOS EM GEOPROCESSAMENTO

A base de dados espacial é o componente de maior custo em um sistema de Geoprocessamento. A abrangência dos trabalhos realizados permite distinguir entre dois tipos de ambientes:

- *projetos de análise espacial* sobre regiões de pequeno e médio porte;
- *inventários espaciais* sobre grandes regiões.

Como exemplo do primeiro caso, pode-se tomar o caso de RIMAs (Relatórios de Impacto Ambiental), como no caso de criação de uma hidroelétrica ou de uma ferrovia, por exemplo. No segundo caso, situam-se levantamentos

sistemáticos, como os feitos pelo INPE para mapear o desflorestamento na Amazônia.

PROJETOS DE ANÁLISE ESPACIAL

Projetos de análise espacial usualmente requerem grande flexibilidade e abrangência das funções do SIG, para quantidade limitada de dados, mas para grande variedade destes. Para ilustrar, tome-se exemplos de aplicações realizadas com o SGI/INPE:

- PETROBRAS/CENPES (estudo da bacia do Recôncavo): 50 "planos de informação" para uma área de cerca de 5.000 km², com diferentes tipos de dados (geofísica, magnetometria, mapa geológico, fotos aéreas), e grande quantidade de dados derivados (conversão de DTM para imagens, transformações IHS->RGB, entre outros).
- EMBRAPA/NMA (zoneamento da ilha de Fernando de Noronha): 80 "planos de informação" para uma área de 3.000 km², com mapas de declividade, uso do solo, geomorfologia, direção dos ventos, localização de espécies, locais acessíveis para turismo, entre outros.

Considera-se que a tecnologia atual de BDG já é capaz de dar suporte a este tipo de aplicações.

PROJETOS DE INVENTÁRIO ESPACIAL

No caso de *inventário*, a ênfase maior é no tratamento de grandes bases de dados, sendo que os mesmos procedimentos são repetidos para todos os dados. As regiões abrangidas são muito grandes, como pode ser visto nos exemplos:

- INPE (Atlas de desmatamento da Amazônia): 228 "projetos", para uma área total de quase 5 milhões de km². Cada projeto possui 5 "planos de informação" (fisionomia florestal, hidrografia, rede viária, base municipal e desflorestamento).

- SOS Mata Atlântica (mapeamento dos remanescentes): 200 projetos, para uma área total de 3 milhões de km². Cada projeto possui 5 "planos de informação".

Nos casos de inventário, o trabalho se resume, em muitos casos, às fases de interpretação dos dados, integração à base de dados espacial e plotagem de resultados.

2.6 REFERÊNCIAS GERAIS SOBRE GEOPROCESSAMENTO

Para uma discussão mais aprofundada sobre Geoprocessamento, o leitor deve procurar Burrough (1986), Star e Estes (1990), ou a coletânea editada por Maguire, Goodchild e Rhind (1991). Em Português, merecem ser citados os trabalhos de Assad e Sano (1993) e Paredes (1995).

3 EVOLUÇÃO DA TECNOLOGIA DE GEOPROCESSAMENTO

“Se você conhece o inimigo e conhece a si mesmo, não precisa temer o resultado de cem batalhas. Se você se conhece mas não conhece o inimigo, para cada vitória sofrerá também uma derrota. Se você não conhece o inimigo nem a si mesmo, perderá todas as batalhas”.

Sun Tzu (A Arte da Guerra, sec. V a.C.).

3.1 APRESENTAÇÃO

Este capítulo traça uma visão retrospectiva e prospectiva sobre a tecnologia de Geoprocessamento, necessária para situar o tema desta tese, onde consideramos a existência de três gerações de sistemas de informação geográfica.

A *primeira geração* (“CAD cartográfico”) caracteriza-se por sistemas herdeiros da tradição de Cartografia, com suporte de bancos de dados limitado e cujo paradigma típico de trabalho é o mapa (chamado de “cobertura” ou de “plano de informação”). Desenvolvidos a partir da início da década de 80 para ambientes da classe VAX e - a partir de 1985 - para sistemas PC/DOS, esta classe de sistemas é utilizada principalmente em projetos isolados, sem a preocupação de gerar arquivos digitais de dados. Esta geração também pode ser caracterizada como sistemas orientados a projeto (“project-oriented GIS”).

A *segunda geração* de SIGs (“banco de dados geográfico”) chegou ao mercado no início da década de 90 e caracteriza-se por ser concebida para uso em ambientes cliente-servidor, acoplado a gerenciadores de bancos de dados relacionais e com pacotes adicionais para processamento de imagens. Desenvolvida em ambientes multiplataforma (UNIX, OS/2, Windows) com interfaces baseadas em janelas, esta geração também pode ser vista como sistemas para suporte à instituições (“enterprise-oriented GIS”).

Pode-se prever, para o final da década de 90, o aparecimento de uma *terceira geração* de SIGs (“bibliotecas geográficas digitais” ou “centros de dados geográficos”), caracterizada pelo gerenciamento de grandes bases de dados geográficos, com acesso através de redes locais e remotas, com interface via WWW (World Wide Web).

Para esta terceira geração, o crescimento dos bancos de dados espaciais e a necessidade de seu compartilhamento com outras instituições requer o recurso a tecnologias como *bancos de dados distribuídos e federativos*. Estes sistemas deverão seguir os requisitos de *interoperabilidade*, de maneira a permitir o acesso de informações espaciais por SIGs distintos.

A terceira geração de SIG pode ainda ser vista como o desenvolvimento de sistemas orientados para troca de informações entre uma instituição e os demais componentes da sociedade (“society-oriented GIS”). A figura 3.1 ilustra a evolução da tecnologia de Geoprocessamento.

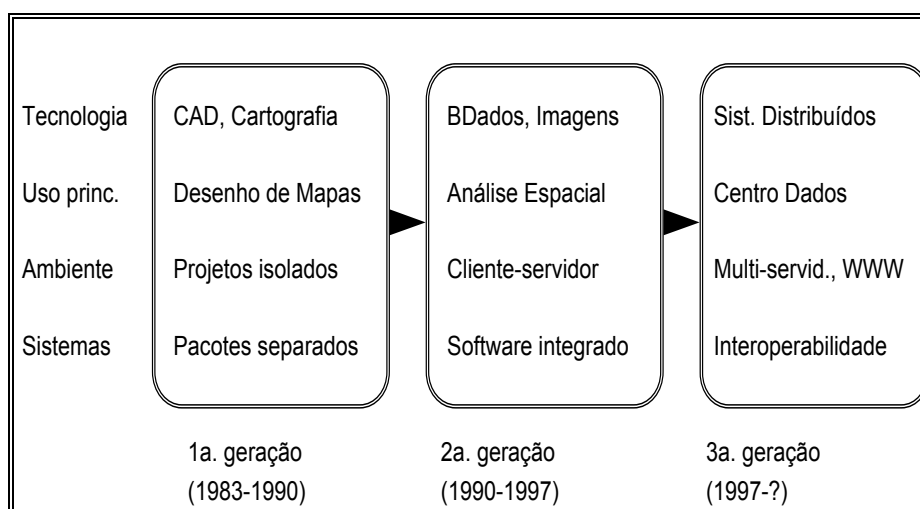


Figura 3.1 - Evolução da tecnologia de SIG.

A evolução da tecnologia e uso de SIG serão detalhadas nas seções seguintes. Na seção 3.2, descrevemos brevemente as características da primeira geração. A seção 3.3 contém uma discussão em maior detalhe sobre as partes constituintes de um SIG de segunda geração (“banco de dados geográfico”). A seção 3.4 inclui uma apresentação dos principais aspectos de uma biblioteca geográfica digital.

3.2 A PRIMEIRA GERAÇÃO: CAD CARTOGRÁFICO

CARACTERIZAÇÃO

A primeira geração de SIG caracteriza-se por sistemas com operações gráficas e de análise espacial sobre arquivos (“flat files”). Sua ligação com gerenciadores de bancos de dados é parcial (parte das informações descritivas se encontra no sistema de arquivos) ou inexistente.

Mais adequados à realização de *projetos de análise espacial* sobre regiões de pequeno e médio porte, estes sistemas enfatizam o aspecto de mapeamento. O sistema permite a entrada de dados sem definição prévia do esquema conceitual, assemelhando-se assim a ambientes de CAD que possuem a capacidade de representar projeções cartográficas e de associar atributos a objetos espaciais. Por força de sua concepção, tais ambientes não possuem suporte adequado para construir grandes bases de dados espaciais.

UTILIZAÇÃO

A primeira e natural utilização dos SIGs na maioria das organizações é como ferramentas para produção de mapeamento básico. Dada a carência de informações adequadas para a tomada de decisões sobre problemas urbanos e ambientais no Brasil, esta maneira de utilizar a tecnologia de Geoprocessamento já representa uma valiosa contribuição. Desafortunadamente, muitas instituições tomam pouco cuidado em manter uma base de dados digital de seus resultados. Em consequência, muitos trabalhos de grande importância estão inacessíveis.

Para citar apenas um exemplo, o projeto “SOS Mata Atlântica” realizou um dos maiores estudos mundiais com a tecnologia de SIG. Foram produzidas mais de 200 cartas na escala 1:250.000, contendo o levantamento de todos os remanescentes da floresta tropical original, a partir da foto-interpretação de imagens de satélite. Apesar do excelente resultado obtido, não houve o cuidado de consolidar os resultados num banco de dados geográficos.

3.3 A SEGUNDA GERAÇÃO: BANCOS DE DADOS GEOGRÁFICOS

CARACTERIZAÇÃO

A segunda geração de sistemas se caracteriza por sistemas concebidos para operar como um *banco de dados geográfico*, entendido como um banco de dados não-convencional aonde os dados tratados possuem, além de atributos descritivos, uma representação geométrica no espaço geográfico. A segunda geração de sistemas ainda está em desenvolvimento, com novos produtos sendo projetados e lançados. Nenhum produto disponível no mercado hoje atende plenamente aos requisitos aqui discutidos, embora os principais fabricantes de sistemas tenham indicado planos em direções compatíveis com as discutidas a seguir.

REQUISITOS

A extensão da tecnologia de bancos de dados para tratar dados geográficos requer avanços em vários pontos, inclusive:

- *Modelagem conceitual*

Avanços na modelagem conceitual em Geoprocessamento são necessários para quebrar a dicotomia matricial-vetorial (NCGIA, 1989) e para gerar interfaces com maior conteúdo semântico (Egenhofer, 1989).

- *Integração Sensoriamento Remoto - Geoprocessamento*

Para aplicações ambientais, um dos requisitos mais importantes para a análise espacial é a integração entre mapas temáticos, modelos de terreno e imagens de satélites (Ehlers et al., 1989). Motivação adicional advém do grande progresso recente na área de Processamento de Imagens de Satélite, com a disponibilidade de novos sensores (como os radares de abertura sintética) e de técnicas de interpretação automática (como a segmentação e redes neurais).

- *Representações Topológicas em Múltiplas Escalas e Projeções*

O gerenciamento de um banco de dados geográfico de grandes dimensões requer que se mantenham múltiplas representações geométricas associadas ao mesmo dado geográfico. Que se pense, por exemplo, nas diferentes representações associadas ao rio Amazonas num banco de dados que cubra toda a região e que esteja particionado geograficamente de acordo com os fusos UTM. Em casos como este, o sistema deve separar o dado de sua representação e garantir a unicidade de descrição.

- *Linguagens de Consulta, Manipulação e Apresentação*

A segunda geração de SIG requer linguagens de consulta, manipulação e representação de objetos espaciais de grande poder expressivo. As linguagens de consulta podem ser baseadas em SQL (como a apresentada em Egenhofer, 1994) e a manipulação de dados geográficos segue a linha da álgebra de mapas (Tomlin, 1990; Câmara, Freitas e Cordeiro, 1994).

- *Novas Técnicas de Análise Geográfica*

A incorporação de novas técnicas de análise geográfica como classificação contínua (Druck e Braga, 1995) e modelagem ambiental (Kemp, 1993) é necessária para que os SIG sejam capazes de satisfazer de forma plena aos requisitos de análise e modelagem de grandes bases de dados espaciais.

- *Arquiteturas para Bancos de Dados de Grande Porte*

A complexidade e o tamanho dos dados geográficos impõem mudanças nos esquemas tradicionais de arquiteturas de sistemas de gerência de banco de dados. Além de avanços em arquiteturas (com o uso de tecnologias como sistemas distribuídos) será preciso dispor de novos métodos de indexação espacial, adequados às massas de dados a ser gerenciadas.

Todos estes requisitos serão abordados no contexto desta tese. No capítulo 4 trataremos dos três primeiros pontos (modelagem conceitual, integração SR-SIG

e representações múltiplas). No capítulo 5 e 6, abordaremos as duas questões seguintes (linguagens e novas técnicas de análise geográfica). No capítulo 7, trataremos do problema de propor novas arquiteturas para sistemas de gerência de bancos de dados geográficos.

Utilização

Apesar dos grandes benefícios dos sistemas de segunda geração para as aplicações de Geoprocessamento, seu uso efetivo ainda é muito restrito no Brasil. As principais razões são de ordem *institucional*.

O uso de ambientes cliente-servidor requer o concurso de competência em administração em Bancos de Dados e em Redes de Computadores, muitas vezes não disponível em instituições usuárias de Geoprocessamento. Também exige um maior investimento para adquirir, instalar e operar sistemas gerenciadores de bancos de dados (SGBD) de mercado. Adicionalmente, o grande potencial integrador de SIG só poderá ser exercido quando o sistema se integra ao processo produtivo. Isto requer que as bases de dados corporativas sejam transformadas para o mesmo ambiente de SGBD utilizado pelo SIG.

Um excelente exemplo do uso de SIG em ambiente cliente-servidor é o sistema SAGRE, desenvolvido pelo CPqD/TELEBRÁS. A partir do suporte oferecido por um SIG (VISION) e por um SGBD com campos longos (ORACLE), foram construídas facilidades para operação e gerenciamento de rede telefônica. Este ambiente está sendo instalado em concessionárias de telefonia em todo o Brasil.

3.4 A TERCEIRA GERAÇÃO: BIBLIOTECAS GEOGRÁFICAS DIGITAIS CARACTERIZAÇÃO

Uma *biblioteca geográfica digital* (ou um “centro de dados geográfico”) é um banco de dados geográfico compartilhado por um conjunto de instituições. Esta biblioteca deve ser acessível remotamente e armazenar, além dos dados geográficos, descrições acerca dos dados (“metadados”) e documentos multimídia associados (texto, fotos, áudio e vídeo).

Este novo paradigma é motivado pelo aguçar da nossa percepção dos problemas ecológicos, urbanos e ambientais, pelo interesse em entender, de forma cada vez mais detalhada, processos de mudança local e global e pela necessidade de compartilhar dados entre instituições e com a sociedade.

O núcleo básico de uma biblioteca geográfica digital é um grande banco de dados geográficos. Para ilustrar, imaginemos dois cenários: um grande centro ambiental brasileiro e uma secretaria de planejamento de uma prefeitura de médio porte (cerca de 1.000.000 de habitantes):

- Cenário 1 (“*Banco de Dados da Amazônia*”): Este banco conteria informações detalhadas sobre a ocupação humana na região, incluindo temas como temas básicos (vegetação, pedologia, geomorfologia), ocupação humana (desmatamento, grandes projetos agropecuários, áreas de prospecção mineral), temas derivados (zoneamentos econômicos) e imagens de satélite atualizadas. Este banco de dados poderia ser mantido por uma instituição central (como o IBAMA) e permitir o acesso de forma concorrente por pesquisadores de todo o país.
- Cenário 2 (“*Prefeitura de Curitiba*”): Este banco conteria todas as informações necessárias para planejamento da cidade, incluindo: lotes, quadras, ruas, equipamentos urbanos (hospitais, escolas), redes de água, esgoto e luz. Poderia ser consultado on-line pelas diversas secretarias municipais, por concessionárias de serviço e por cidadãos.

Nos dois cenários, o ambiente deve garantir acesso concorrente a uma comunidade de usuários, com diferentes métodos de seleção, incluindo folheamento ("browsing") e linguagem de consulta.

Um exemplo desta tecnologia é o sistema EOSDIS (EOS Data and Information System), projetado pela NASA para atender ao programa EOS ("Earth Observation System"), que inclui um conjunto de satélites com enorme capacidade de geração de dados (2 Tb/dia).

Tendo em vista o panorama geral, as próximas seções apresentam os principais requisitos da nova geração de SIG, divididos em grandes áreas: *disponibilidade de metadados, acesso via WWW, navegação pictórica ("browsing") e interoperabilidade.*

REQUISITOS: METADADOS

A idéia de metadados (ou "dados sobre os dados") é criar um ambiente que apresente descrições gerais sobre os conjuntos de dados disponíveis localmente ou em centros associados. Exemplo é o "Master File Directory" (MFD) da NASA, que inclui referências para conjuntos de dados em diferentes instituições e agências espaciais em outros países.

Um dos principais desafios com a construção de ambientes de metadados é balancear o esforço requerido para descrever as coleções de dados, pois a informação final deve ser suficiente para guiar a busca (Abel and O'Callaghan, 1992). O MFD da NASA adota a filosofia de prover um conjunto mínimo de descritores obrigatórios, com campos em texto livre para informação adicional; esta estratégia minimiza tanto o esforço requerido para compor o relatório quanto a capacidade de busca disponível.

Outro aspecto desejável em ambientes de metadados é a disponibilidade de dados síntese, na forma de mapas em escala muito reduzida que possam ser utilizados para localizar geograficamente os conjuntos de dados disponíveis.

O ideal é permitir um refinamento do processo de consulta, estabelecendo um caminho contínuo entre o nível mais abstrato de metadados e os dados (Smith

and Frank, 1989). Assim, a partir de informação de carácter geral (“Banco de Dados da Amazônia”), passamos a visões de carácter regional (“Estado de Roraima”), específico (“Mapas de Vegetação de Roraima”) até chegar no dado propriamente dito (“Mapa de Vegetação da Região do Rio Demene em Roraima na escala 1:100.000”).

REQUISITOS: ACESSO VIA WWW

A disponibilidade de interfaces multimídia via Internet, proporcionada pelo ambiente WWW, permite que os dados geográficos sejam apresentados de forma pictórica (através de mapas reduzidos e imagens “quick-look”).

No entanto, o ambiente WWW (com o uso de HTML - Hyper Text Mark-up Language) apresenta alguns problemas para uso com bancos de dados geográficos, principalmente porque a linguagem HTML é navegacional e não suporta a noção de transação².

Para poder combinar de forma apropriada a tecnologia WWW/HTML com um ambiente de consulta típico de bancos de dados geográficos, é inevitável lançar mão, além dos recursos pictóricos do HTML, de uma linguagem de consulta com restrições espaciais.

REQUISITOS: NAVEGAÇÃO PICTÓRICA

A navegação pictórica (*browsing*) pode ser visto como uma *seleção baseada em apontamento*: uma interface interativa permite ao usuário percorrer o banco de dados.

Prover mecanismos de navegação é importante, pois não se pode supor que o usuário saiba *a priori* quais os tipos de dados disponíveis e como fazer para ter acesso a estes. As ferramentas de navegação permitem ao usuário ter acesso ao dados com base em sua *localização espacial*. A figura 3.2 mostra um possível particionamento por mapas-índice de um país, que poderia ser utilizado no processo de navegação.

²Esta observação, bem como muitas considerações importantes sobre o problema de bibliotecas geográficas digitais, são devidas a Danton Nunes (gerente do nó GRID-INPE), a quem o autor agradece.

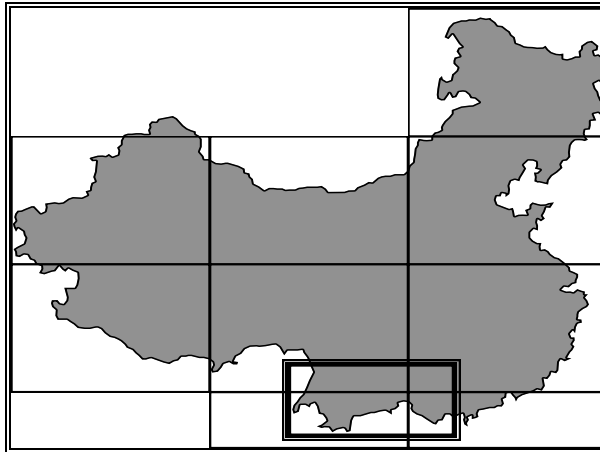


Figura 3.2 - Estrutura de mapas-índice para navegação.

O grande desafio ao montar um ambiente de navegação é garantir rapidez de resposta e interatividade. Para garantir rapidez, é necessário dispor de mecanismos de generalização, que devem ser diferentes conforme os formatos de dados subjacentes. Pode ser necessário, por razões de desempenho, pre-computar os dados necessários ao processo de folheamento.

REQUISITOS: INTEROPERABILIDADE

O compartilhamento de dados e procedimentos entre bancos de dados geográficos baseados em SIGs distintos é um desafio considerável. Como não existe um modelo de dados geográficos universalmente aceito (ao contrário do modelo relacional para aplicações convencionais), os diversos SIG do mercado apresentam diferenças significativas na maneira de operar e nos formatos internos de armazenamento. Para tentar remediar estes problemas, vários países tem estabelecido padrões cartográficos de transferência de dados, que almejam preservar a riqueza da informação geográfica (topologia e atributos). Estes padrões buscam uma “neutralidade” com relação aos diferentes fabricantes, e incluem o SDTS (Spatial Data Transfer Standard) nos EUA, o SAIF (Spatial Archive and Interchange Format) no Canadá e o OGIS (Open GIS Standard)³. Na realidade, estes padrões foram projetados para atuar em níveis distintos, como descrito a seguir.

³O autor agradece a Andrea Hemerly pelas informações sobre o SAIF, SDTS e OGIS.

O SDTS é um padrão de *transferência de dados espaciais*, baseado num modelo de dados que distingue entre *entidades espaciais* (entidades do mundo real), *objetos espaciais* (elementos gráficos simples, com ou sem topologia) e fenômenos espaciais (relacionamentos entre entidades e objetos). A semântica do modelo de dados SDTS inclui ainda informações sobre qualidade de dados. Informações sobre SDTS podem ser obtidas em <ftp://sdts.er.usgs.gov>.

O SAIF é uma linguagem de *especificação e troca de dados*, baseada num modelo de dados orientado-a-objeto que procura modelar os diferentes tipos de dados geográficos, estabelecendo entre eles relações de identidade, generalização e agregação. Informações sobre o SAIF podem ser obtidas no endereço WWW <http://www.wimsey.com/infosafe/saif>.

O OGIS é um *padrão de interoperabilidade entre SIGs*, baseado em tecnologia de objetos distribuídos (CORBA e OLE), que prescreve um modelo de alto nível de programação de aplicações em SIG (APM - “Applications Programming Model”). O modelo APM foi concebido para facilitar a implementação transportável de aplicações e serviços. O modelo de objetos geográficos do OGIS é baseado no modelo SAIF, o que permitirá estabelecer uma compatibilidade entre estes dois padrões.

No entanto, mesmo que se estabeleça uma transferência de dados entre bancos de dados geográficos com diferentes formatos, deve-se lembrar que, em aplicações complexas como Geoprocessamento, parte substancial da informação está contida nos procedimentos de consulta, manipulação e apresentação⁴. Assim, o intercâmbio de dados não garante a interoperabilidade. No contexto desta tese, não temos condição de explorar em maior detalhe este assunto fascinante, que esperamos poder tratar posteriormente

⁴Para citar um exemplo nacional, o sistema SAGRE (da TELEBRÁS) consiste em 200.000 linhas de aplicações para redes telefônicas desenvolvidas na linguagem do sistema GEOVision.

3.5 CONTRIBUIÇÃO DO TRABALHO NA EVOLUÇÃO DA TECNOLOGIA

Com base no panorama de evolução da tecnologia de Geoprocessamento apresentado neste capítulo, esta tese pretende atuar em duas perspectivas:

- estabelecer sólidas bases conceituais para o desenvolvimento e o aprimoramento de um SIG de segunda geração (SPRING);
- lançar idéias para o projeto e desenvolvimento da tecnologia de centros de dados geográficos, que virá para complementar as capacidades do SPRING.

Nosso objetivo é projetar um SIG que satisfaça a todos às necessidades de um sistema de segunda geração, os requisitos descritos na seção 3.3, “Bancos de Dados Geográficos” serão abordados nos capítulos que seguem em maior detalhe. Acreditamos que a existência de um banco de dados geográfico completo é fundamental para que possamos construir bibliotecas geográficas digitais eficientes.

4 MODELAGEM DE DADOS EM GEOPROCESSAMENTO

“Na realidade, a ciência e a arte vêm aos homens por intermédio da experiência, porque a experiência criou a arte e a inexperiência, o acaso. E a arte aparece quando, de um conjunto de noções experimentadas, se exprime um único juízo universal de casos semelhantes”.

Aristóteles (Metafísica).

4.1 APRESENTAÇÃO

Um *modelo de dados* é um conjunto de ferramentas conceituais utilizado para estruturar dados num sistema computacional. Aspecto fundamental no projeto de um SIG, o modelo descreve como a realidade geográfica será representada no computador. Nenhuma outra decisão limita tanto a abrangência e o crescimento futuro do sistema quanto a escolha do modelo de dados.

A necessidade de modelos gerais que representem avanços conceituais em Geoprocessamento foi assim expressa no programa de pesquisa do NCGIA (1989):

" Current GIS are either raster-based or vector-based and have limited abilities to combine the two. (...) While current systems are built using one model or the other, an ideal GIS would offer both and allow full interaction between them".

Como indicado por Goodchild (1992a), a indústria de Geoprocessamento amadureceu a tal ponto que questões sobre estruturas de dados, algoritmos e funcionalidade contam com respostas padronizadas. Nesta situação, a modelagem de dados tem um papel crítico ao determinar a capacidade de uso e a rapidez de aprendizado do sistema.

Este capítulo apresenta um modelo de dados para Geoprocessamento, projetado com técnicas de orientação-por-objetos e que permite lidar com os diversos tipos de dados ambientais. As contribuições mais relevantes do modelo são:

- Apresentar uma abordagem unificada para a modelagem em Geoprocessamento, combinando as idéias de campos e objetos.
- Integrar imagens de Sensoriamento Remoto e Modelos Numéricos de Terreno com mapas temáticos, mapas cadastrais e redes.
- Prover suporte para representações geométricas múltiplas de uma mesma entidade do mundo real e permitir a coexistência de representações vetorial, matricial e grades num mesmo sistema.
- Engendrar uma interface de alto nível com conteúdo semântico.

Este modelo está sendo utilizado como base no projeto e implementação do sistema SPRING (Câmara et al., 1992) em sua versão 2.0. No âmbito do SPRING, o modelo serve como base para a definição de interface com o usuário, e de uma biblioteca de classes para SIGs em C++.

4.2 TRABALHOS ANTERIORES NA LITERATURA

VISÃO GERAL

O trabalho mais influente neste assunto é o de Goodchild (1992b). Nele, o autor apresenta as noções de *campos* e *objetos* como representações conceituais para dados geográficos. A maior parte dos trabalhos relativos a questão de modelos de dados em Geoprocessamento utiliza o paradigma de orientação por objetos, e serão descritos aqui em prioridade.

MODELAGEM ORIENTADA-A-OBJETOS EM GEOPROCESSAMENTO

O uso de conceitos de orientação por objetos em Geoprocessamento foi analisado pela primeira vez em Egenhofer e Frank (1987). Neste trabalho, os autores mostram como aplicar conceitos de *classificação*, *especialização*, *generalização* e *agregação* ao projeto de SIGs. Para uma revisão recente dos resultados teóricos em uso de orientação por objetos em projeto de SIG, veja-se Egenhofer (1995).

Uma das principais linhas de pesquisa deriva da comunidade de bancos de dados, ao procurar estender os conceitos de bancos de dados orientados a objetos

ao domínio geográfico. Merecem destaque os trabalhos de Worboys (Worboys et al., 1990; Worboys, 1994), do grupo ligado ao SGBD O₂ (Scholl and Voisard, 1991; David and Voisard, 1993; Rigaux and Scholl, 1995) e de grupos brasileiros (Times e Salgado, 1995).

Os modelos propostos pela comunidade de banco de dados tendem a analisar um banco de dados geográfico como uma coleção de entidades individuais (objetos). Exemplo pode ser visto no trabalho de David e Voisard (1993):

“A map is a collection of geographic objects. A geographic object is usually derived from an entity of the real world. It has a spatial part as well as an alphanumeric description.”

Esta visão confunde o uso de modelos de dados orientados-a-objetos com a decomposição do espaço em objetos discretos. Veja-se este trecho do trabalho de Worboys (1994):

“There are two broad and opposing classes of models of geographical information. The class of field-based models treats such informations as collections of spatial distributions. (...) Patterns of topographic altitudes, rainfall and temperature fit this view. The class of entity-based models treats the information space as populated by discrete, identifiable entities (or objects), each with a geo-reference. (...) These two models result in two opposed GIS implementation approaches, namely vector and raster. Object-oriented models decompose the space into objects”.

Particularmente interessante é o conceito de “realm”, introduzido por Güting e Schneider (1993). Um *realm* descreve a geometria do espaço geográfico utilizando uma base discreta (conjuntos de coordenadas inteiras). Esta idéia é utilizada para desenvolver um álgebra formal (Güting et al., 1995) que define operações sobre o espaço discreto de um *realm*.

O modelo de dados orientado-a-objetos mais completo disponível é o proposto para o padrão SAIF (Spatial Archive and Interchange Format), utilizado no Canadá como formato de arquivo e intercâmbio de dados geográficos (Survey

and Resources Mapping Branch, 1993; Sondheim, 1993). O padrão SAIF está sendo utilizado como base para implementar as extensões espaciais da linguagem SQL 3.

HISTÓRICO DOS TRABALHOS NO INPE

O modelo de dados proposto reflete a experiência do sistema SGI/INPE. Desenvolvido a partir de 1984, o SGI teve grande aceitação no mercado nacional para aplicações ambientais. O modelo do SGI foi desenvolvido com base nos conceitos de abstração de dados e contemplava a noção de *plano de informação* (Câmara, Velasco e Oliveira, 1982). Tratava-se essencialmente de modelo adequado a análise temática, que embutia a noção de geo-campos.

Uma noção ainda preliminar de *geo-objetos* foi introduzida em Erthal, Alves e Câmara (1988), que foi a base de implementação da primeira versão do SPRING. Posteriormente, Alves (1990) apresentou um modelo de dados orientado-a-objetos que apresentava avanços conceituais importantes. As discussões que levaram à proposta deste modelo foram conduzidas num grupo aonde participaram Marco Casanova, Andrea Hemerly, Ubirajara Freitas, Ricardo Cartaxo, João Argemiro Paiva e o autor. As noções de *categoria* e de *visual* são devidas a Ricardo Cartaxo e a de *banco de dados geográficos*, a Ubirajara Freitas. Coube ao autor formalizar os conceitos, explicitar os diferentes níveis conceituais, introduzir a noção de *geo-objetos e objetos cadastrais* e elaborar a forma final do modelo. O autor expressa ainda seu débito a Diógenes Alves pelas idéias expressas em Alves (1990) e por ter-lhe indicado o trabalho de Goodchild (1992b).

COMPARAÇÃO COM TRABALHOS DA LITERATURA

O modelo de dados aqui apresentado distingue-se dos trabalhos anteriores por sua visão abrangente das classes de dados presentes em SIGs e na separação entre os universo conceitual e o universo de representação. No nível conceitual, o modelo é uma evolução das idéias de Goodchild (1992b) e seu universo de representação é compatível com o padrão SAIF.

4.3 VISÃO GERAL DO PROCESSO DE MODELAGEM

O processo de Modelagem é a forma que dispomos para traduzir o mundo real em outros domínios. Um das abordagens mais úteis para este problema é o chamado “paradigma dos quatro universos” (Gomes e Velho, 1995), que distingue:

- O universo do *mundo real*, que inclui as entidades da realidade a ser modeladas no sistema.
- O universo *matemático (conceitual)*, que inclui uma definição matemática (formal) das entidades a ser incluídas no modelo.
- O universo de *representação*, onde as diversas entidades formais são mapeadas para representações geométricas.
- O universo de *implementação*, aonde as estruturas de dados e algoritmos são escolhidos, baseados em considerações como desempenho, capacidade do equipamento e tamanho da massa de dados. É neste nível que acontece a codificação.

Esta visão do processo de modelagem está ilustrada na figura 4.1.

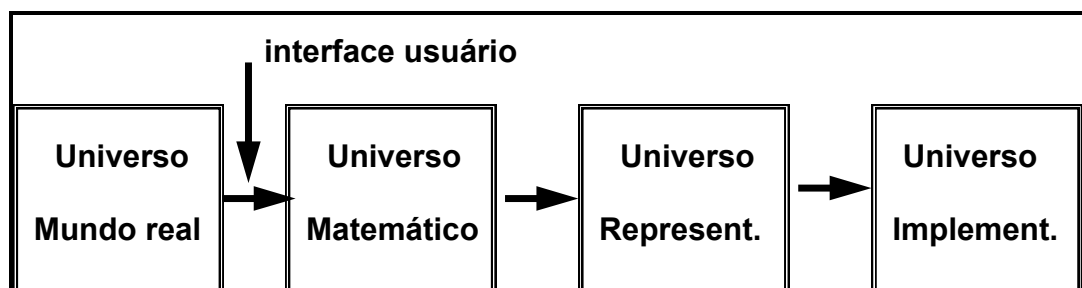


Figura 4.1 - Modelagem segundo o paradigma dos quatro universos.

A visão apresentada não se limita a sistemas de Geoprocessamento, mas representa uma perspectiva unificadora aos problemas de Computação Gráfica e Processamento de Imagens. Sua aplicação ao problema de Geoprocessamento é particularmente apropriada pois permite equacionar os problemas da área, como se pode constatar:

- no universo do *mundo real*, encontramos os fenômenos a ser representados (tipos de solo, cadastro urbano e rural, dados geofísicos e topográficos).

- no universo *conceitual* (*matemático*), pode-se distinguir entre as grandes classes formais de dados geográficos (*dados contínuos* e *objetos individualizáveis*), e especializar estas classes nos tipos de dados geográficos utilizados comumente (dados temáticos e cadastrais, modelos numéricos de terreno, dados de sensoriamento remoto).
- no universo de *representação*, as entidades formais definidas no universo conceitual são associadas a diferentes representações geométricas, que podem variar conforme a escala e a projeção cartográfica escolhida e a época de aquisição do dado. Aqui se distingue entre as representações matricial e vetorial, que podem ainda ser especializadas.
- o universo de *implementação* é onde ocorre a realização do modelo de dados através de linguagens de programação. Neste universo, escolhemos estruturas de dados (tais como árvores quaternárias e árvores-R) para implementar as geometrias do universo de representação.

Com base nesta visão, as dicotomias tradicionais de Geoprocessamento (campos-objetos e matricial-vetorial) podem ser resolvidas, mostrando-se que elas encontram-se *em níveis distintos de abstração*.

Esta análise também indica que a interface de usuário de um SIG deve, tanto quanto possível, refletir o universo conceitual e esconder detalhes dos universos de representação e implementação. No nível conceitual, o usuário lida com conceitos mais próximos de sua realidade, e minimiza a complexidade envolvida nos diferentes tipos de representação geométrica⁵.

⁵No entender do autor, este representa um dos principais diferenciais do software SPRING com relação a seus concorrentes.

4.4 O UNIVERSO DO MUNDO REAL

Pretende-se construir um modelo não limitado a uma particular área de aplicação de Geoprocessamento e que possa incorporar aplicações tão diversas como Estudos Ambientais, Agricultura, Geologia e Redes.

Para tanto, será preciso mapear as diversas áreas de aplicação de Geoprocessamento, em função dos tipos de dados utilizados, como ilustrado na Tabela 4.1. Como se verá a seguir, o modelo proposto é aplicável a todas as áreas listadas, ilustrativas da diversidade de problemas em Geoprocessamento.

Tabela 4.1

Aplicações típicas de Geoprocessamento

<i>Aplic.</i>	<i>Escalas típicas</i>	<i>Tipos de dados</i>	<i>Represent. Gráficas</i>	<i>Operações</i>
Floresta	1:10.000 a 1:500.000	cadastro rural (talhões)	matricial, vetorial	classif. imagens, consulta espacial
Agricult.	1:25.000 a 1:250.000	dados temáticos, S. Remoto, MNT	matricial, vetori algrades, TIN	análise espacial, declividade
Geologia	1:100.000 a 1:2,500,000	MNT, dados temáticos, imag ens	matricial, grades	transf. IHS, visualiz. 3D
Redes	1:1.000 a 1:10.000	redes lineares (topologia)	vetorial	consulta espacial, cálculos dedicados
Estudos Urbanos	1:1.000 a 1:25.000	redes, cadastro urbano	vetorial	consulta espacial

4.5 O UNIVERSO CONCEITUAL

VISÃO GERAL

Da análise do universo do mundo real pode-se constatar que as aplicações de Geoprocessamento lidam com dois grandes tipos de dados (Goodchild, 1992b):

- *Geo-Campos*: correspondem a grandezas distribuídas espacialmente, como tipo de solo, topografia e teor de minerais.
- *Objetos geográficos (ou geo-objetos)*: individualizáveis e tem identificação com elementos do mundo real, como lotes num cadastro urbano e postes numa rede elétrica. Estes objetos tem atributos não-espaciais e podem estar associados a mais de uma representação geométrica.

O aspecto fundamental da distinção entre campos e geo-objetos é a questão da identidade. Existem milhares de áreas no Brasil classificáveis como “latossolo roxo”, mas há apenas um “Jardim Botânico do Rio de Janeiro”. Os *geo-campos* obedecem ainda ao princípio do “*planar enforcement*” (restrição de preenchimento do plano): cada localização de um *geo-campo* está associada a um e somente um valor da variável espacial representada. No caso de *geo-objetos*, o princípio do “*planar enforcement*” não precisa ser obedecido, pois os objetos podem ser disjuntos no espaço ou o mesmo local do espaço corresponder a mais de um objeto. Tome-se, a título de exemplo, os casos de solos e de rede elétrica. Num geo-campo de solos, cada local do espaço está associado a um tipo específico de solo (p.ex. Latossolo Roxo); numa de rede elétrica, num mesma localização podem conviver geo-objetos distintos (um poste com um transformador e uma chave).

Nossa abordagem segue a filosofia do modelo de dados do SGBD orientado-a-objetos O₂ (Lécluse, Richard e Velez, 1991), que suporta os conceitos de *identidade de objetos*, *tipos de dados* e *herança*, e considera os seguintes conjuntos disjuntos:

- O conjunto *I* de identificadores de dados geográficos.

- O conjunto C de nomes de classes geográficas.
- O universo U de atributos *descritivos* $\{A_1, \dots, A_n\}$, definidos nos domínios $D(A_1), \dots, D(A_n)$. Denotaremos por a_i um valor possível do atributo A_i .

Seguiremos os seguintes passos:

1. Definir as classes básicas do modelo e estabelecer as relações entre elas, dentro dos princípios de especialização, generalização e agregação.
2. Estabelecer como, a partir do modelo, é possível definir um esquema conceitual para um banco de dados geográfico, por especialização das classes básicas.

CLASSES DO UNIVERSO CONCEITUAL

As classes básicas do modelo, definidas a seguir, são: *geo-campo*, *geo-objeto*, *objeto cadastral*, *objetos não-espaciais*, *plano de informação* e *banco de dados geográfico*.

Definições Preliminares

Inicialmente, será importante estabelecer a base geométrica na qual as classes do modelo são definidas. Seguindo a idéia de Güting e Schneider (1993), utilizamos um conjunto *discreto* de pontos como suporte para as entidades espaciais.

Definição 4.1. Projeção geográfica planar.

Dadas uma região da superfície terrestre definida por suas coordenadas de latitude e longitude $\varphi = (\varphi_0, \varphi_f)$, $\lambda = (\lambda_0, \lambda_f)$ e uma transformação cartográfica $\rho: (\vartheta, \Lambda) \rightarrow \mathfrak{R}^2$ que mapeia as coordenadas terrestres (φ, λ) para pontos no \mathfrak{R}^2 , uma *projeção geográfica planar* G induzida por φ, λ e ρ é uma região do \mathfrak{R}^2 definida como:

$$(x_0, y_0) = \rho(\varphi_0, \lambda_0), \quad (y_f, x_f) = \rho(\varphi_f, \lambda_f)$$

$$G = \{ (x, y) \in \mathfrak{R}^2 \mid x_0 \leq x \leq x_f, \quad y_0 \leq y \leq y_f \}.$$

A partir de uma região contínua da superfície terrestre, podemos definir nosso conceito de *região geográfica* (ou *reticulado geográfico*).

Definição 4.2. Região Geográfica.

Dada uma projeção planar G , uma escala χ associada a esta projeção e $\Delta x, \Delta y \in \mathbb{R}$, uma *região geográfica* $R = [G, \chi, \Delta x, \Delta y]$ é um reticulado conectado contido no plano \mathbb{Z}^2 tal que $R = \{ (x_i, y_j) \in G \mid x_i = i \cdot \Delta x, y_j = j \cdot \Delta y \text{ para } i, j \in \mathbb{Z} \}$

A região geográfica R serve de suporte geométrico para localizações geográficas, pois toda uma localização geográfica será representada por um ponto de R . O uso de um conjunto discreto de pontos facilita uma definição formal das classes de dados geográficas e das operações associadas. A definição de região geográfica proposta não restringe a escolha da representação geométrica (matricial ou vetorial) associada aos objetos geográficos.

Sabemos, da Cartografia, que uma área contínua da superfície da Terra será mapeada para uma superfície plana no \mathbb{R}^2 através de transformações projetivas, para uma projeção cartográfica e escala definidas. A teoria cartográfica também nos ensina (Joly, 1990) que todo mapa possui uma precisão associada, que é usualmente definida em função de sua escala. Com base nestas considerações, associaremos as dimensões $\Delta x, \Delta y$ do reticulado R à precisão cartográfica necessária para representar univocamente cada elemento geográfico, dadas uma área da superfície terrestre, uma escala e uma projeção. Deste modo, mesmo se posteriormente (no universo de representação), escolhermos uma geometria vetorial para objetos definidos no universo conceitual, o reticulado R ainda poderá ser utilizado como base geométrica.

Geo-Campos

Definição 4.3. Geo-Campo.

Um geo-campo é um objeto $f = [R, V, \lambda]$, onde R é uma região geográfica, V um domínio de valores e $\lambda: R \rightarrow V$ um mapeamento entre pontos em R e valores em V .

Um geo-campo representa a distribuição espacial de uma variável que possui valores em todos os pontos de uma região geográfica. Os geo-campos podem ser especializados em:

- TEMÁTICO : a função λ caracteriza um *geo-campo temático* quando define um mapeamento $\lambda: R \rightarrow V$ tal que V é um conjunto finito enumerável. Os elementos de V definem os *temas* de um geo-campo (p.ex. um geo-campo de pedologia é caracterizado pelo conjunto de temas {latosolo roxo, litosolo, cambisol, ...}).
- NUMÉRICO - uma instância desta classe, chamada de geo-campo numérico (ou MODELO NUMÉRICO DE TERRENO) é definida através de uma função $\lambda: R \rightarrow V$ tal que V é o conjunto dos reais (p. ex. um geo-campo de altimetria).
- DADO_SENSOR_REMOTO - esta classe é uma especialização de NUMÉRICO, definida a através de uma função $\lambda: R \rightarrow V$, tal que V é o conjunto dos naturais. Na prática, uma instância desta classe é obtida através de discretização da resposta recebida por um sensor (passivo ou ativo) de uma área da superfície terrestre.

Denotamos a classe de geo-campos por GEOCAMPO⁶. No caso de geo-campo temático, cada tema pode ser composto de um ou mais qualificadores⁷ e pode ter atributos. Por exemplo, num geo-campo de solos, cada tema de solo ({latosolo roxo, litosolo, cambisol, ...}) pode estar associado a uma tabela de atributos (como teor de sódio, nitrogênio e potássio, pH do solo, umidade). Note-se que um dado de sensor remoto é um caso particular de geo-campo numérico, onde o valor é proporcional à resposta do solo medida por um sensor numa faixa do espectro eletromagnético.

As figuras abaixo apresentam exemplos de geo-campos.

⁶Na notação utilizada neste trabalho, o uso de *SMALLCAPS* denota uma classe.

⁷O caso mais usual é um mapa temático onde cada localização está associada a um tema único. No entanto, é muito útil poder definir uma situação aonde mais de um conjunto de temas está associado a cada localização.



Figura 4.2 - Dado Sensor Remoto (Imagem LANDSAT de Manaus).

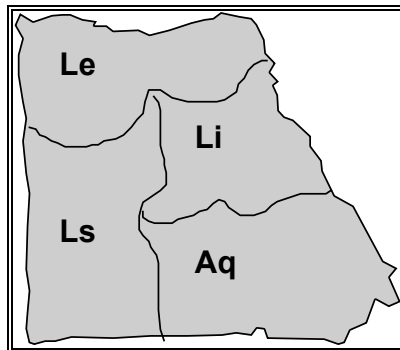


Figura 4.3 - Geo-campo temático (solos)

Especializações de Geo-Campos

As classes básicas TEMÁTICO, MNT e DADO_SENSOR_REMOTO podem ser utilizadas para derivar uma hierarquia de especialização, que reflita o domínio de aplicação. Por exemplo, para estudos de uso do solo, poderíamos criar uma classe TIPO_DE_SOLO (especialização de TEMÁTICO), cujas instâncias são os geo-campos de uso do solo da região de estudo, e também criar uma classe IMAGEM_LANDSAT (especialização de DADO_SENSOR_REMOTO) cujas instâncias seriam imagens do satélite LANDSAT sobre a região de estudo.

Geo-Objeto

Definição 4.4. Geo-Objeto

Dados um conjunto de regiões geográficas R_1, \dots, R_n e um conjunto de atributos A_1, \dots, A_n com domínios $D(A_1), \dots, D(A_n)$, um *geo-objeto* é um objeto $go = [a_1, \dots, a_n, S_1, \dots, S_m]$, composto pelos valores a_i (aonde $a_i \in D(A_i)$) e por localizações geográficas S_i (aonde $S_i \subseteq R_i$). Denotamos o valor do i -ésimo atributo descritivo de go por $go.A_i$ e a i -ésima localização geográfica de go por $go.R_i$.

Em outras palavras, um geo-objeto é um elemento único, que possui atributos descritivos e está associado a múltiplas localizações geográficas. Esta definição tem três grandes motivações adicionais:

1. *As projeções cartográficas:* A projeção planar da Terra, a partir de escalas macroregionais, é feita com o uso de quadrículas que estão particionados em sistemas de referência independentes, que definem recortes arbitrários no espaço e podem dividir a localização de um geo-objeto. Por exemplo, um particionamento cartográfico da Amazônia na projeção UTM, escala 1:250.000 faz com que os principais rios tenham representações geométricas descontínuas em vários mapas.
2. *Representações geométricas em diferentes escalas:* Na prática, num mesmo banco de dados geográfico, podem conviver representações da mesma realidade geográfica em diferentes escalas geográficas. Por exemplo, consideremos um conjunto de mapas dos municípios do Estado de São Paulo, que inclui um mapa geral (na escala de 1:1.000.000) e mapas regionais (na escala de 1:500.000). Nesta situação, um mesmo geo-objeto (p.ex., o município de São José dos Campos) teria duas representações geométricas: uma no mapa geral e outra no mapa regional do Vale do Paraíba.
3. *Múltiplas representações temporais:* As diferentes representações de um mesmo objeto podem corresponder a variações temporais do mesmo, como no caso de um lago que teve suas bordas alteradas.

Para ilustrar este conceito, consideremos a figura 4.4, que mostra um banco de dados da Amazônia, onde os quadrados pontilhados representam o recorte espacial do banco de dados. Entidades como os rios Amazonas e Xingu têm representações em diferentes particionamentos espaciais do banco de dados.

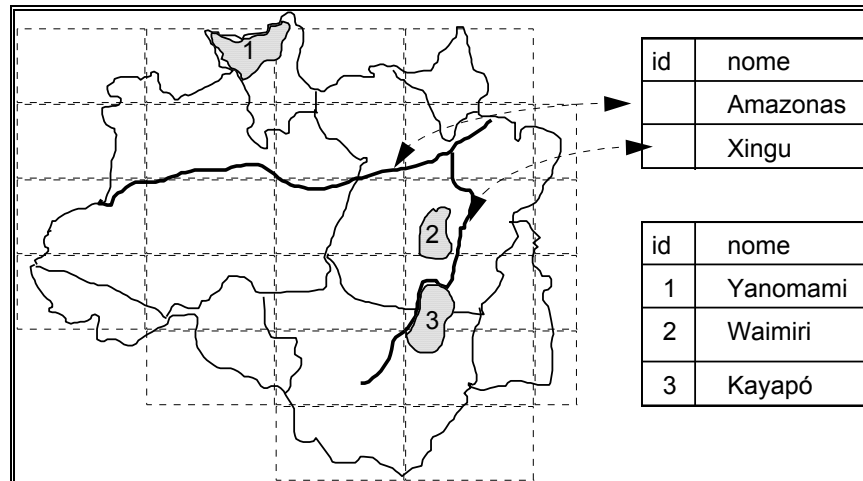


Figura 4.4 - Princípio de identidade em bancos de dados geográficos.

Objeto não-espacial

Em muitas situações, é conveniente permitir a associação de informações não-espaciais a um banco de dados geo-referenciado. Por exemplo, considere-se uma aplicação de cadastro urbano em uma prefeitura, que já dispõe de um sistema para cálculo do IPTU baseado num cadastro alfanumérico de lotes. Neste caso, pode-se desejar associar o cadastro alfanumérico a dados georeferenciados contendo a localização geográfica destas propriedades. Para englobar estas entidades, introduzimos a noção de objeto não-espacial.

Definição 4.5. Objeto não-espacial

Dado um conjunto de atributos A_1, \dots, A_n , um objeto não-espacial é um objeto $on = [a_1, \dots, a_n]$ definido pelos atributos $a_i \in D(A_i)$, e que não possui localizações espaciais associadas.

Assim, a noção de objeto não-espacial engloba qualquer tipo de informação que não seja georeferenciada e que se queira agregar a um SIG. Denotamos a classe de objeto não-espaciais por NÃOESPACIAL.

O exemplo anexo mostra o caso de uma aplicação de cadastro rural. Neste caso, temos os geo-objetos da classe “fazendas” (que estão referenciados espacialmente) e desejamos estabelecer a ligação entre estes geo-objetos e a informação alfanumérica já existente, sob a forma de um cadastro de propriedades. Neste caso, as informações de cadastro são consideradas um *objeto não-espacial*.

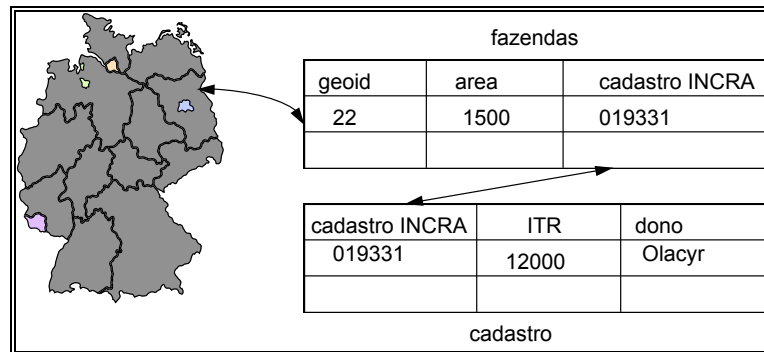


Figura 4.5 - Exemplo de ligação entre geo-objeto e objeto não-espacial.

Objeto Cadastral (Complexo)

Nossa definição de geo-objeto permite a associação de objetos geográficos a diferentes regiões no espaço. Como as aplicações de Geoprocessamento usualmente não armazenam ou manipulam elementos isolados, é conveniente armazenar a representação geométrica de um geo-objeto em conjunto com seus vizinhos, mantendo as relações de topologia. Por exemplo, num cadastro urbano, os lotes de um mesmo bairro são armazenados e apresentados em conjunto. Esta característica nos leva a introduzir a idéia de *objeto cadastral*, um *objeto complexo* que agrupa geo-objetos para uma dada projeção cartográfica e região geográfica, sendo instâncias da classe CADASTRAL.

Definição 4.6. Objeto Cadastral.

Um objeto cadastral é um objeto $oc=[R,GO,geo]$, onde R é uma região geográfica, GO é um conjunto de geo-objetos e $geo: GO \rightarrow 2^R$ é um mapeamento que atribui a cada geo-objeto $go \in GO$, um subconjunto não-vazio de 2^R (ou seja, uma localização no mapa).

A relação entre o objeto cadastral e os geo-objetos nele presentes é *ESTÁ_MAPEADO_EM*, definida pelo mapeamento *geo(go)*. A classe dos objetos cadastrais será denotada por *CADASTRAL*. A figura 4.6 ilustra um exemplo de objeto cadastral, para o caso de um cadastro urbano. Cada objetos cadastral deste classe contém os mapeamentos dos geo-objetos das classes *LOTES*, *HOSPITAIS* e *ESCOLAS*.

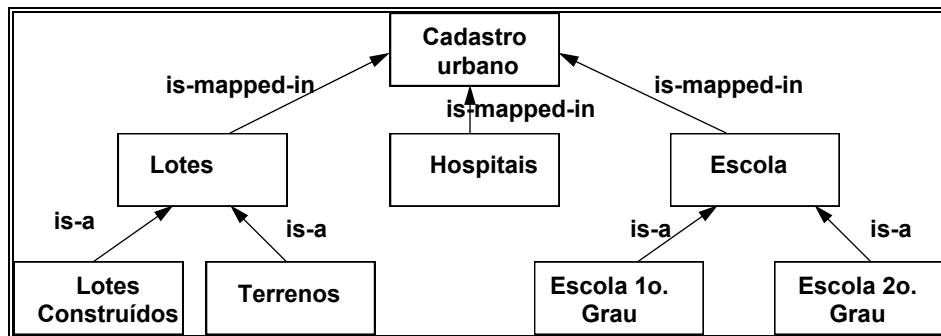


Figura 4.6 - Exemplo de objetos complexos (cadastro urbano).

Para ilustrar o conceito e o uso de objetos cadastrais, consideramos a ainda a figura 4.7, onde temos um caso de um objeto cadastral (cadastro urbano) contendo os lotes de uma parte do bairro “Jardim Esplanada” em São José dos Campos. Os geo-objetos da classe “lotes” incluem todos os lotes existentes na cidades (na figura apresentados sob forma de tabela). Como o obejto cadastral está limitado espacialmente, contém apenas as representações geométricas de uma parte dos geo-objetos (na figura, indicados pelas setas).

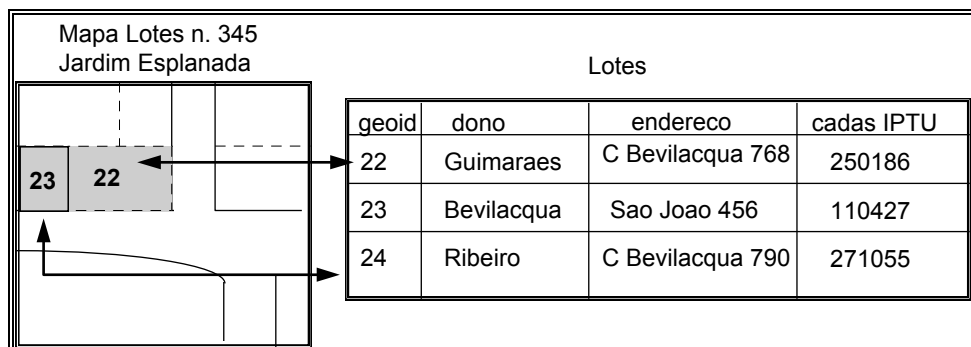


Figura 4.7 - Ilustração do conceito de objeto cadastral.

Consideramos ainda a especialização da classe *CADASTRAL* na classe *REDE*. Uma instância desta classe é um objeto $oc=[R,GO,geo]$, onde o mapeamento

$geo:GO \rightarrow 2^R$ é tal que $geo(go)$ é um ponto ou uma linha em R . Como exemplo desta classe de objetos, considere-se a especificação da classe REDE na classe REDE ELÉTRICA. Uma instância desta classe conterá os mapeamentos de diferentes classes de geo-objetos, como mostrado na figura 4.8.

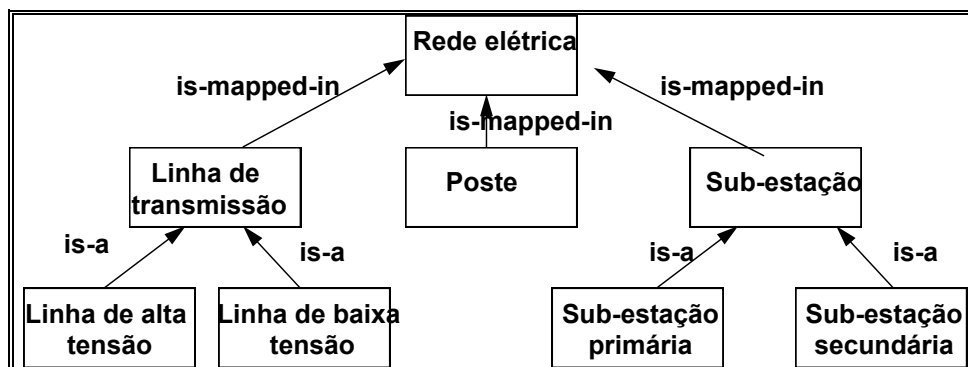


Figura 4.8 - Exemplo de objeto complexo (rede elétrica).

A figura 4.9 mostra o caso de uma rede elétrica, representada através de uma topologia arco-nó. Note-se que uma mesma localização geográfica está associada a dois geo-objetos (um poste e um transformador).

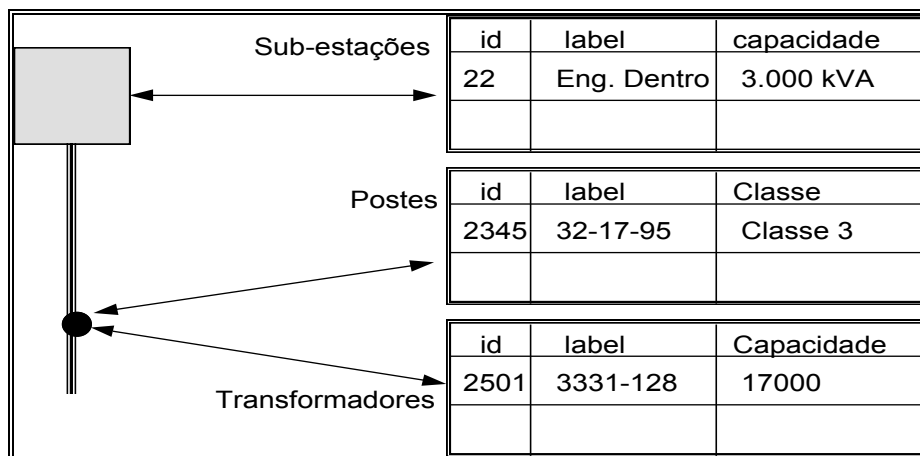


Figura 4.9 - Ilustração do conceito de rede.

Plano de Informação

Um conceito muito útil, para fins de definição de interface e de operações, é a idéia de plano de informação. Trata-se da generalização dos conceitos de geo-campos e de objeto cadastral.

Definição 4.7. Plano de informação

Um plano de informação é um objeto cadastral ou um geo-campo.

Uma instância da classe PLANO DE INFORMAÇÃO representa, para uma dada região geográfica, o lugar geométrico de um conjunto de dados geográficos (um geo-campo ou um objeto cadastral).

Banco de dados geográficos**Definição 4.8. Banco de dados geográficos**

Um banco de dados geográficos é um objeto $bdg = [P, GO, ON]$, composto por um conjunto de planos de informação $P = \{pi_1, pi_2, \dots, pi_n\}$, um de geo-objetos $GO = \{go_1, go_2, \dots, go_n\}$ e um de objetos não-espaciais $ON = \{on_1, on_2, \dots, on_n\}$.

Esta definição é particularmente interessante pois não faz restrição sobre escala dos dados e nem sobre a continuidade espacial dos planos de informação que compõem o banco. Tome-se, por exemplo, um banco de dados geográficos sobre a Amazônia Legal, com os seguintes dados:

- Vegetação (fito-fisionomia), especialização da classe TEMÁTICO, com 780 temas de vegetação definidas pelo IBGE (projeto RADAM): 26 cartas digitalizadas a partir de mapas na escala 1:1.000.000, projeção Lambert.
- Imagens de satélite, especialização da classe DADO_SENSOR_REMOTO, obtidas pelo satélite LANDSAT TM, em três anos (1986, 1990 e 1994): 228 imagens, projeção UTM.
- Geo-campos de uso do solo, especializações da classe TEMÁTICO, com as classes: floresta primária, floresta secundária, floresta de várzea, cerrado, área de regeneração, rios e nuvem. Os mapas finais serão compatíveis com a escala 1:250.000, projeção UTM.
- Zoneamento econômico-ecológico do estado de Roraima, especialização da classe CADASTRAL, com 15 planos de informação gerados a partir do estudos na escala 1:100.000.

Resumo do universo CONCEITUAL

Em resumo, o universo de representação do modelo tem como entidade básica um *banco de dados geográfico*. Este banco é composto por *planos de informação*, por *geo-objetos* e por *objetos não-espaciais*. Os planos de informação podem ser *geo-campos* ou *objetos cadastrais*. Estas classes podem ser ainda especializadas para construir o esquema conceitual do banco de dados geográfico. A figura 4.10 ilustra estas relações.

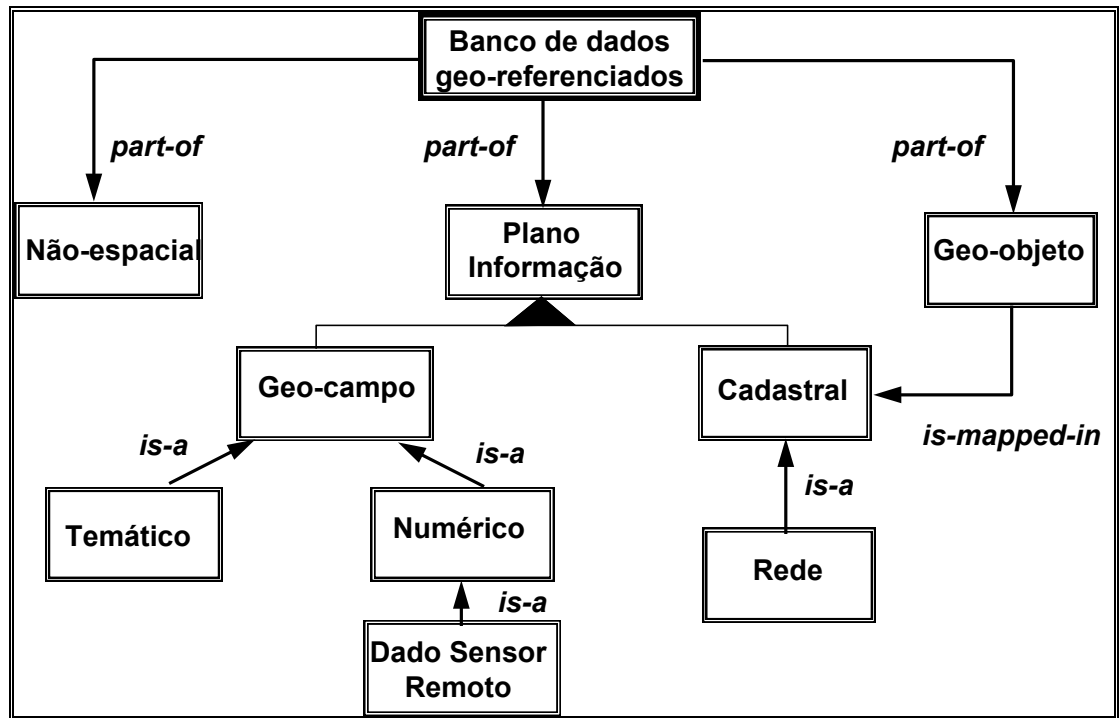


Figura 4.10 - Universo conceitual.

4.6 UNIVERSO DE REPRESENTAÇÃO

No universo de representação, estaremos definindo as possíveis representações geométricas que podem estar associadas às classes do universo conceitual. Inicialmente, devemos considerar as duas grandes classes de representações geométricas: REPRESENTAÇÃO VETORIAL e REPRESENTAÇÃO MATRICIAL. Para detalhar o universo de representação, adotamos proposta semelhante à apresentada pelo padrão SAIF (Sondheim, 1993). A estrutura aqui proposta é uma simplificação do modelo do SAIF, e aproxima-se bastante da implementação utilizada no SPRING-2.0.

Hierarquia de representações matriciais

A REPRESENTAÇÃO GEOMÉTRICA MATRICIAL pode ser especializada segundo a hierarquia de classes mostrada na figura 4.11.

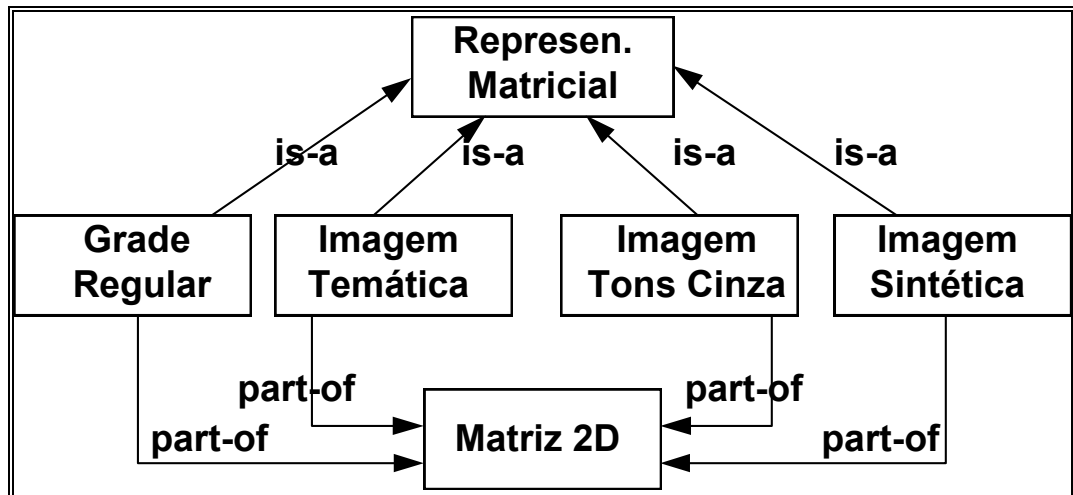


Figura 4.11 - Hierarquia de classes para REPRESENTAÇÃO MATRICIAL.

As classes derivadas de REPRESENTAÇÃO MATRICIAL são:

- **GRADE REGULAR:** uma grade regular é uma matriz de reais.
- **IMAGEM EM TONS DE CINZA:** imagem representada através de uma matriz sobre Q , onde $Q \subset \mathbb{N}$ denota o conjunto dos níveis de quantização da imagem.
- **IMAGEM TEMÁTICA:** representação matricial de um geo-campo TEMÁTICO, dada por uma tripla (M, V, T) , onde M é uma matriz sobre $Q \subset \mathbb{N}$, V são os valores possíveis do geo-campo, e $T: Q \rightarrow V$ uma função injetiva que associa valores em Q a temas em V . Por exemplo, numa imagem temática, um elemento da matriz de valor 2 pode estar associado ao tema “Floresta Ombrófila”.
- **IMAGEM SINTÉTICA:** representação de uma imagem em cores, dada por um par (M, T) , onde M é uma matriz sobre $Q \subset \mathbb{N}$ e por T uma função que associa, para cada valor $q \in Q$, uma tripla de valores inteiros (r, g, b) que definem uma composição colorida do tipo RGB (vermelho, verde,

azul). Utilizada para mostrar imagens em composição colorida em placas gráficas falsa-cor.

Hierarquia de representações vetoriais

Para definir esta hierarquia, precisamos inicialmente precisar melhor o que entendemos pelas primitivas geométricas: *coordenadas 2D*, *coordenadas 3D*, *nó 2D*, *nó rede*, *arcos*, *arcos orientados*, *isolinhas* e *polígonos* (A definição de polígono foi obtida em Puppo e Dettori, 1995).

Dada uma região geográfica R definida como

$$R = \{ (x_i, y_j) \subset [x_0, x_f] \times [y_0, y_f] ; x_i = i \cdot \Delta x, y_j = j \cdot \Delta y \quad i, j \in \mathbb{Z}, \Delta x, \Delta y \in \mathbb{R} \},$$

podemos definir:

- COORDENADA2D: Uma *coordenada 2D* é um objeto composto por uma localização singular $(x_i, y_j) \in R$.
- COORDENADA3D: Uma *coordenada 3D* é um objeto composto por uma localização singular (x_i, y_j, z) , onde $(x_i, y_j) \in R$.
- PONTO2D: Um *ponto 2D* é um objeto que possui atributos descritivos e uma *coordenada 2D*.
- LINHA2D: Uma *linha 2D* possui atributos e inclui um conjunto de *coordenadas 2D* $\{ (x_1, y_1), \dots, (x_n, y_n) ; (x_i, y_i) \in R, x_{i+1} = x_i \text{ ou } y_{i+1} = y_i \}$. Denotaremos a primeira coordenada de uma linha l por $l(0)$, a última por $l(1)$ e o conjunto das demais por l^0 .
- ISOLINHA: uma *isolinha* contém uma *linha 2D*, associada a um valor real (cota).
- ARCO ORIENTADO: um arco orientado contém uma *linha 2D*, associada a uma orientação de percorrimeto.
- NÓ2D: um *nó 2D* inclui uma *coordenada 2D* $(x_i, y_i) \in R$ e uma lista L de linhas 2D, tal que $\forall l_i, l_j \in L (l_i \neq l_j), (x_i, y_i) = l_i(1) = l_j(0)$. Trata-se da

conexão entre duas ou mais linhas, utilizada para manter a topologia da estrutura;

- NÓ REDE: um *nó de rede* contém um *nó 2d* e uma lista de *arcos orientados*, onde a cada instância associamos uma impedância e um custo de percorrimeto.
- NÓ 3D: uma instância desta classe contém uma *coordenada 3D* (x_i, y_i, z) e um lista L de *linhas 2D*, tal que $\forall l_i, l_j \in L (l_i \neq l_j), (x_i, y_i) = l_i(1) = l_j(0)$. Trata-se da conexão entre três ou mais linhas de uma grade triangular.
- POLÍGONO: Para caracterizar a classe dos polígonos, precisamos inicialmente de uma definição auxiliar, a de *área*. Uma *área* é um objeto a formado por um conjunto de linhas 2D, tal que $a = \{ l_1, \dots, l_n \}$ e $l_{i+1}(0) = l_i(1), i \in [1, n-1]$ e $l_1(0) = l_n(1)$. Uma área a divide o reticulado R em um conjunto *interior*, denotado por a^0 , e um conjunto *exterior*, denotado por \hat{a} . Dado uma área a_e (“área externa”) $\in R$, um conjunto de áreas internas (“buracos”) $\{ a_1, \dots, a_n \}$ tais que $a_i \subseteq a_e (i=1, \dots, n)$, um *polígono* é um objeto definido pelo conjunto de pontos $pg = \{ a_e^0 \cap \hat{a}_1 \cap \hat{a}_2 \dots \cap \hat{a}_n, i=1, \dots, n \}$. Neste contexto, um polígono contém uma lista de *linhas 2D* e uma lista de *nós 2D*, que descrevem as coordenadas da área externa e das áreas internas que compõem o polígono.

Uma vez definidas as primitivas geométricas vetoriais, podemos estabelecer a hierarquia de representações geométricas vetoriais, como mostrado na figura 4.12. Na figura 4.12, distinguimos os relacionamentos de especialização (“*is-a*”), inclusão de uma instância (“*part-of*”), inclusão de um conjunto de instâncias (“*set-of*”) e inclusão de uma lista de identificadores de instâncias (“*list-of*”). Este último relacionamento será utilizado para manter a topologia 2D.

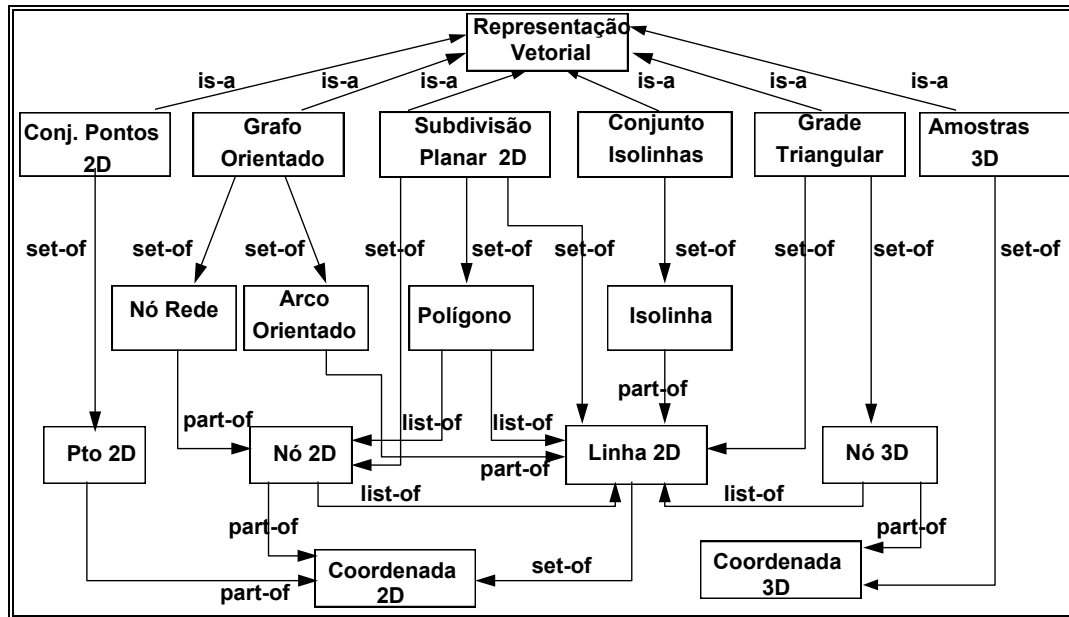


Figura 4.12 - Hierarquia de classes para REPRESENTAÇÃO VETORIAL.

. Consideramos aqui, as seguintes especializações de REPRESENTAÇÃO VETORIAL:

- CONJUNTO DE PONTOS 2D: uma instância desta classe é um conjunto de *pontos 2d*, utilizado para guardar localizações isolados no espaço (p.ex. no caso de poços de petróleo).
- CONJUNTO DE ISOLINHAS: uma instância desta classe é um conjunto I de *isolinhas* e $\forall i_i, i_j \in I (i_i \neq i_j), i_i \cap i_j = \emptyset$. Num mapa de isolinhas, cada linha possui uma cota e as linhas não se intersectam.
- SUBDIVISÃO PLANAR: dada uma região geográfica R , uma instância desta classe contém um conjunto Pg de *polígonos*, L de *linhas 2D* e N de *nós 2D*.
- GRAFO ORIENTADO: uma instância desta classe é uma representação, composta de um conjunto de *nó rede* e de um conjunto de *arco orientado 2d*.
- GRADE TRIANGULAR: uma instância desta classe contém um conjunto de *nós 3D* e um conjunto L de *linhas 2D* tal que:

$$\forall l_i \in L, \exists l_j, l_k, l_m, l_n \in L (i \neq j \neq k \neq m \neq n), \text{ onde}$$

$$l_i \cap l_j = l_i(0) = l_i \cap l_k \text{ e } l_i \cap l_m = l_i(1) = l_i \cap l_n \text{ e}$$

$$l_i^0 \cap l_p = \emptyset, \forall l_p \in L, i \neq p.$$

Trata-se de uma grade onde todas as linhas se interseptam, mas apenas em seus pontos iniciais e finais.

- MAPA PONTOS3D: uma instância desta classe é um conjunto de *coordenadas 3d*. Trata-se de um conjunto de amostras 3D.

Relacionamento entre plano de informação e representação geométrica

No universo conceitual, definimos que uma generalização para as classes de objetos que possuíam suporte espacial, através da classe PLANO DE INFORMAÇÃO. Para relacionar esta classe com suas possíveis representações, introduzimos a classe REPRESENTAÇÃO GEOMÉTRICA como generalização de REPRESENTAÇÃO MATRICIAL e de REPRESENTAÇÃO VETORIAL, e o relacionamento *está_representado_em* para denotar a relação entre as classes PLANO DE INFORMAÇÃO e REPRESENTAÇÃO GEOMÉTRICA. A figura 4.13 indica a hierarquia básica do universo de representação.

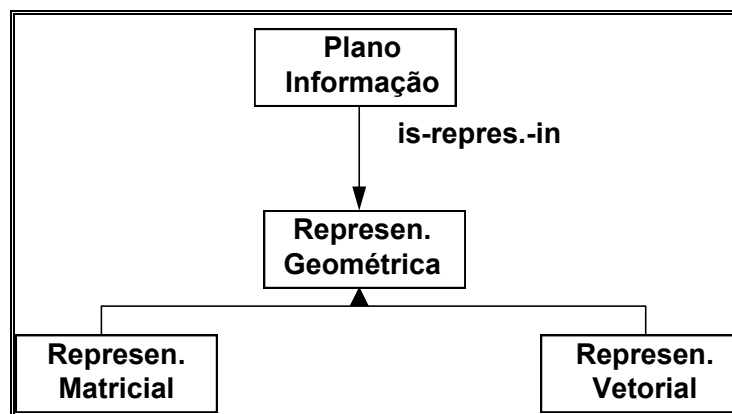


Figura 4.13 - Relacionamento entre universos conceitual e de representação.

Para maiores detalhes sobre as representações geométricas associadas a cada classe do universo conceitual, veja-se a seção 4.8 (“Relação entre os universos do modelo”).

4.7 UNIVERSO DE IMPLEMENTAÇÃO

Ao discutir o universo de implementação, estaremos indicando quais as estruturas de dados a ser utilizadas para construir um sistema de Geoprocessamento. Aqui tratamos com decisões concretas de programação e que podem admitir número muito grande de variações. Estas decisões podem levar em conta as aplicações às quais o sistema é voltado, a disponibilidade de algoritmos para tratamento de dados geográficos e o desempenho do hardware. Para uma discussão sobre os problemas de implementação de operações geográficas, veja-se Güting et al. (1995).

Um dos aspectos principais a ser levados em conta no universo de implementação é o uso de estruturas de indexação espacial. Os métodos de acesso a dados espaciais compõem-se de *estruturas de dados e algoritmos de pesquisa e recuperação* e representam um componente determinante no desempenho total do sistema. Apanhados gerais da literatura são feitos em Cox Junior (1991) e Rezende (1992). Estes métodos operam sobre chaves multidimensionais e dividem-se conforme a representação dos dados associados: pontos (ex: *árvores K-D*), linhas e polígonos (ex: *árvores R e R+*) e imagens (ex: *árvores quaternárias*). Estes e outros métodos têm possibilitado (principalmente no caso de pontos e linhas) grandes melhorias de desempenho no acesso a dados geográficos.

O grande fator limitante da maior parte dos métodos de acesso estudados é que foram projetados para operar em *memória principal*. Num SGBDG de grande porte, é preciso fazer acesso eficiente a dados em *memória secundária*. Isto vale tanto para dados vetoriais como dados matriciais. No caso de dados vetoriais, Mediano, Casanova e Dreux (1994) apresentam uma nova proposta de utilizar uma extensão das *árvores-B* para poder apresentar apenas os dados geográficos relevantes para uma determinada escala, sem incorrer em procedimentos desnecessários. Esta estrutura, chamada de *árvore-V*, permite ainda o acesso a dados multiresolução e é muito conveniente como suporte a métodos de folheamento de um grande banco de dados.

4.8 RELAÇÃO ENTRE OS UNIVERSOS DO MODELO

O paradigma dos “quatro universos de modelagem” (Gomes e Velho, 1995) parte do princípio que o mapeamento entre cada universo não é reversível e admite alternativas. Discutiremos a seguir estas relações.

Do mundo real para o universo conceitual

A passagem do mundo real para o universo conceitual pode admitir algumas variações, conforme o domínio de aplicação. Em alguns casos, o mapeamento é direto: por exemplo, as imagens de satélite e grandezas topográficas e geofísicas são naturalmente mapeadas para instâncias de GEOCAMPO. No caso de mapas municipais e de divisão política, sua associação com as classes GEOOBJETO e MAPA DE GEOOBJETOS é também direta.

Os levantamentos temáticos podem se prestar a duas interpretações, conforme seu uso: quando se tratar de trabalhos de inventário (como o mapa de vegetação da Amazônia), devem ser modelados como instâncias de GEOCAMPO (ou, mais especificamente, da classe TEMÁTICO). No caso de estudos detalhados em médias e grandes escalas (como no zoneamento ecológico-econômico), onde cada região é caracterizada por qualificadores específicos, é conveniente que estes levantamentos sejam associados a instâncias de GEOOBJETOS e de MAPA DE GEOOBJETOS.

Do universo conceitual para a representação

Este mapeamento apresenta várias alternativas não-excludentes, a saber:

- Instâncias da classe DADO_SENSORE_REMOTO são usualmente armazenadas como IMAGENS EM TONS DE CINZA e IMAGENS SINTÉTICAS;
- Um geo-campo NUMÉRICO pode ser representados tanto como matrizes (GRADE REGULAR) como vetores (CONJUNTO DE ISOLINHAS, GRADE TRIANGULAR e CONJUNTO DE PONTOS 3D);
- um geo-campo TEMÁTICO pode ser representado tanto como vetores topologicamente estruturados (SUBDIVISÃO PLANAR), como por matrizes (IMAGEM TEMÁTICA).

- Instâncias de CADASTRAL são representados por instâncias de SUBDIVISÃO PLANAR ou CONJUNTOS DE PONTOS 2D.
- Uma REDE é representada por um GRAFO ORIENTADO.

A literatura tem consagrado a conclusão de que um SIG de propósito geral deve prover todas as alternativas de representação.

Das representações à implementação

Conforme dissemos anteriormente, a realização do universo de implementação é decisão concreta de programação. Faremos aqui algumas considerações de ordem prática:

- o armazenamento de pontos 3D em árvores K-D (Bentley, 1975) traz um ganho muito significativo para aplicações como a geração de grade regular a partir de um conjunto de amostras esparsas;
- o uso de árvores quaternárias (Samet, 1990) para armazenar imagens em tons de cinza não é eficaz. Para o caso de mapas temáticos, apesar de utilizado em pelo menos um sistema comercial (SPANS), os ganhos não são significativos;
- o uso de árvores-R (Gutman, 1984) ou árvores-V (Mediano et al., 1994) só se torna eficiente quando completado por algoritmos de busca e processamento que utilizam suas propriedades.

Resumo

Para compreender melhor a relação entre os diferentes universos (níveis) do modelo, a tabela a seguir arrola vários exemplos de entidades do mundo real e suas correspondentes no modelo.

TABELA 4.2

CORRESPONDÊNCIA ENTRE UNIVERSOS DO MODELO

<i>Universo do mundo real</i>	<i>Universo conceitual</i>	<i>Universo de representação</i>	<i>Universo de implementação</i>
Mapa de vegetação	Geo-campo Temático	Imagem temática Subdivisão Planar	Matriz 2D Linhas 2D (com R-Tree)
Mapa altimétrico	Geo-campo Numérico	Grade regular Grade triangular Conjunto Pontos 3D Conjunto Isolinhas	Matriz 2D Linhas 2D e Nós 3D Pontos 3D (KD-tree) Linhas 2D
Lotes urbanos	Geo-objetos		
Mapa de lotes	Cadastral	Subdivisão Planar	Linhas 2D (com R-Tree)
Rede elétrica	Rede	Grafo Orientado	Linhas 2D (com R-Tree)

4.9 DEFINIÇÃO DO ESQUEMA DO BANCO DE DADOS GEOGRÁFICOS

O processo de definir o esquema conceitual de um banco de dados geográficos consiste de:

- estender a hierarquia de especialização definida pelo modelo, criando classes derivadas de GEOOBJETO, CADASTRAL, REDE, TEMÁTICO, MODELO NUMÉRICO DE TERRENO e DADO SENSOR REMOTO.
- definir restrições de integridade indicando, para cada nova classe definida como especialização de GEOOBJETO, quais especializações da classe CADASTRAL podem conter descrições de geo-objetos desta classe;

- definir restrições de integridade indicando, para cada especialização de PLANO DE INFORMAÇÃO (GEOCAMPO ou CADASTRAL), quais especializações de REPRESENTAÇÃO GEOMÉTRICA estão associadas.

A última operação só será realizada se o mapeamento entre o nível de representação e o de implementação não for fixo, o que seria uma forma desejável de extensibilidade.

Como exemplo, considere-se a seguinte definição de esquema conceitual para um banco de dados geográficos para cadastro rural (ilustrada na figura 4.14):

- uma classe FAZENDAS, especialização de GEOOBJETO, que pode ainda ser sub-especializada em LATIFÚNDIO e MINIFÚNDIO .
- uma classe MAPA DE PROPRIEDADES, especialização de CADASTRAL, que define um mapeamento para os objetos da classe FAZENDAS e suas especializações.
- uma classe MAPA DE SOLOS, especialização de TEMÁTICO, cujas instâncias armazenam o tipo de solos para as áreas de estudo;
- as classes ALTIMETRIA e DECLIVIDADE, especializações de MODELO NUMÉRICO DE TERRENO, cujas instâncias guardam (respectivamente) a topografia e a declividade da área de estudo;
- uma classe DADOS LANDSAT, especialização de DADO SENSOR REMOTO, cujas instâncias contêm as imagens do satélite LANDSAT sobre a região de estudo.

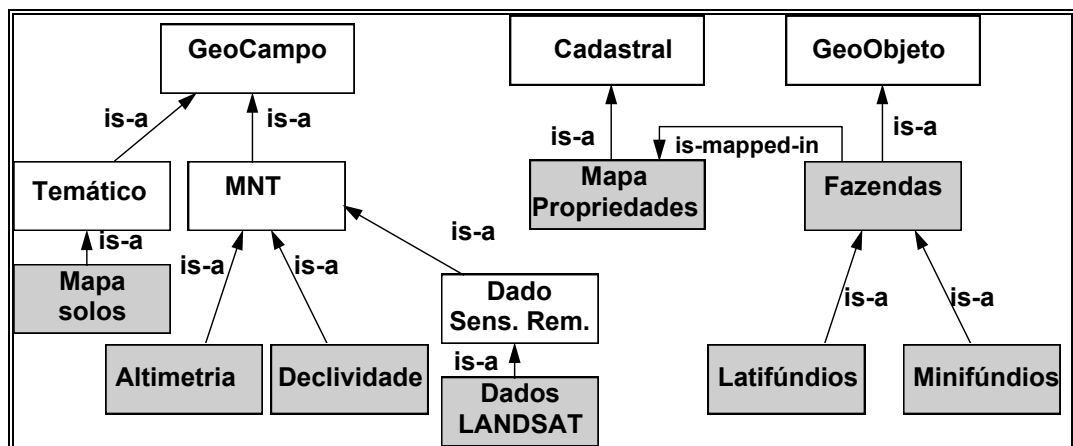


Figura 4.14 - Exemplo de definição de Esquema Conceitual.

4.10 ATRIBUTOS BÁSICOS DAS CLASSES DO UNIVERSO CONCEITUAL

Associamos a cada uma das classes do universo conceitual, atributos básicos que serão utilizados posteriormente pelos métodos definidos para cada classe. Estes atributos incluem (mudar isto.....)

- Para a classe GEOCAMPO, o identificador, nome, resolução horizontal, resolução vertical, escala e projeção cartográfica.
- Para a classe TEMÁTICO, a lista dos *temas* associados.
- Para a classe MNT, os valores máximo e mínimo da função.
- Para a classe GEOOBJETO, o identificador e o nome.

Para a classe CADASTRAL, o identificador, o nome, a escala e a projeção cartográfica.

4.11 INTERFACE COM O USUÁRIO

A nosso ver, um dos aspectos mais interessantes do modelo é permitir o projeto de interfaces amigáveis e eficientes, o que acelera em muito o aprendizado do sistema e aumenta a produtividade do usuário.

No SPRING-2.0, o uso do modelo permitiu o projeto de interface que permite a manipulação de dados geográficos num nível abstrato. Ao selecionar um *plano de informação*, apenas as operações disponíveis para a classe de dados geográficos associada ao plano estarão disponíveis ao usuário. Esta abordagem reduz muito a incerteza na escolha das funções válidas num SIG.

5 OPERAÇÕES SOBRE CAMPOS E OBJETOS GEOGRÁFICOS

“Se quisermos apreender a essência de um complexo de noções abstratas, devemos por um lado investigar as relações mútuas entre os conceitos e as afirmações feitas a seu respeito e, por outro, investigar como eles se relacionam com as experiências.”

Albert Einstein

5.1 INTRODUÇÃO

As operações de consulta e manipulação de dados geográficos constituem a essência de um SIG, ao diferenciar o Geoprocessamento de tecnologias como Cartografia Automatizada e Projeto Auxiliado por Computador. Apesar de sua importância, consideramos que o avanço da tecnologia de Geoprocessamento está a requerer um trabalho conceitual semelhante ao realizado por Codd na área de bancos de dados convencionais, ao definir a álgebra relacional e suas propriedades.

A não-existência de uma álgebra geográfica é um fator limitante no avanço da tecnologia e dos sistemas para Geoprocessamento. Neste capítulo, procuramos dar uma contribuição para o assunto. A partir do modelo de dados apresentado no capítulo anterior, estabelecemos aqui uma taxonomia para as diversas operações de análise geográfica, que serão divididas em: *operadores sobre geo-objetos*, *operadores sobre geo-campos*, *operadores de transformação entre geo-campos e geo-objetos* e *operadores mistos entre geo-objetos e geo-campos*. Esta análise permitirá:

- Obter um entendimento formal sobre a natureza das operações em Geoprocessamento.
- Definir uma linguagem de manipulação e consulta para dados geográficos de propósito geral e ampla aplicação.
- Fornecer uma base conceitual para abordar o problema de otimizar operações em bancos de dados geográficos.

Uma visão completa sobre as operações de SIGs inclui ainda: *operadores de apresentação gráfica de dados e sobre representações geométricas*. Estes operadores não serão discutidos do ponto de vista conceitual, mas apresentados quando da proposta de uma linguagem no capítulo seguinte.

5.2 BREVE REVISÃO DA LITERATURA

A literatura especializada apresenta um grande número de artigos sistematizando as operações em um SIG, como Goodchild (1987), Maguire and Dangerramond (1991) e Burrough (1992). Não obstante, a caracterização adotada neste trabalho é inédita, pois é a primeira derivada da explícita diferenciação entre geo-objetos e geo-campos. Isto tem conseqüências importantes, pois permite compreender melhor a natureza de cada operador.

Os trabalhos da literatura abordam o problema de duas perspectivas distintas: operações de consulta sobre geo-objetos (Egenhofer, 1994) e operações de manipulação sobre geo-campos (Tomlin, 1990), sem unificar as duas visões.

Verifica-se ainda um grande esforço na literatura para formalizar as relações espaciais entre objetos discretos (Clementini et al., 1993; Egenhofer et al., 1994), que tomaremos como base para nossa análise.

Em resumo, nosso trabalho procurar traçar uma ponte entre as duas visões tradicionais de operações geográficas (baseadas em entidades e baseadas em campos), sempre buscando uma visão integradora dos processos de análise espacial. Como tal, temos a expectativa de trilhar caminho inovador no assunto.

5.3 PROPRIEDADES DE ÁLGEBRAS GEOGRÁFICAS

Em sua resenha, Güting (1994) indica algumas propriedades desejáveis de álgebras geográficas. No que segue, apresentamos as principais considerações contidas nesta referência e, em seguida, discute-se como a caracterização proposta as responde.

- **Extensibilidade.** *“A definição de tipos e, em especial, operações, é dependente da aplicação. Deve ser possível definir posteriormente tipos adicionais ou alternativos, o que conduz ao requisito de extensibilidade”.*

Na abordagem proposta, considera-se que o modelo de dados definido no capítulo anterior contém todas as classes básicas para Geoprocessamento. Cada tipo específico de dados (e.g. Uso do Solo) pode ser derivado de um tipo primitivo do modelo (no caso, Temático).

- **Quais tipos de dados?** *“Será realmente necessário ter vários tipos distintos, para distinguir, por exemplo, pontos, linhas e regiões? Uma vantagem de um tipo único é que será mais fácil garantir o fechamento sobre as operações. Por outro lado, vários tipos dão maior expressividade e permitem uma aplicação mais precisa das operações”.*

Um tipo único de dados não permitiria dar conta nem da dualidade básica (campos x objetos), nem das diferentes representações gráficas. É necessário uma hierarquia básica, tanto no universo formal, como no universo de representação.

- **Operações sobre conjuntos.** *“Uma álgebra espacial deveria oferecer não apenas operações sobre tipos espaciais atômicos (como regiões), mas também dados complexos (como um temático)”.*

A proposição feita procura atender aos requisitos dos usuários de SIG e inclui operações sobre conjuntos complexos, tais como mapas de geo-objetos.

- **Completeza.** *“a questão é saber se existe um critério formal para dizer se uma particular coleção de operações é completa de algum modo. Algum sucesso limitado já foi obtido no estudo de operadores topológicos”.*

Dada a variedade de usos de Geoprocessamento, o critério utilizado neste trabalho foi baseado na prática: analisamos as operações disponíveis nos principais SIGs do mercado (e.g. ARC/INFO, SPANS, MGE, SPRING) e levamos em conta a experiência prática do INPE e de seus parceiros no uso de sistemas de informação geográfica (veja-se, por exemplo, Assad e Sano, 1993). Procuramos apresentar um conjunto abrangente, no qual todas as classes de operações deste SIGs estão presentes.

5.4 OPERAÇÕES SOBRE GEO-CAMPOS

Descrevemos nesta seção as operações sobre GEO-CAMPOS e suas especializações TEMÁTICO, NUMÉRICO e DADO_SENSOR_REMOTO, que podem ser classificados como (Tomlin, 1990):

1. *Pontuais*: a saída da operação é um geo-campo cujos valores são função apenas dos valores dos geo-campos de entrada em cada localização correspondente. Podem operar apenas sobre um campo (e.g. *fatiar* um modelo numérico de terreno, *classificar* uma imagem) ou realizar intersecções entre conjuntos espaciais (e.g. *operações booleanas* entre mapas temáticos)⁸.
2. *vizinhança*: o resultado é um geo-campo cujos valores dependem da vizinhança da localização considerada. Exemplos são a *filtragem* espacial de uma imagem e o cálculo de *declividade* de um MNT.
3. *zonais*: caso especial de (2), estas operações são definidas sobre regiões específicas de um geo-campo de entrada, onde as restrições são fornecidas por outro geo-campo temático. Um exemplo seria: “dada um mapa de solos e um mapa de declividade da mesma região, obtenha a declividade média para cada tipo de solo”.
4. *propriedades*: operações que devolvem valores (escalares ou vetores) de um geo-campo ou um conjunto de geo-campos. Exemplos: “calcule a altitude média do terreno” e “obtenha o histograma de uma imagem”.

Operações pontuais

Definição 5.1. Operações pontuais sobre geo-campos.

Seja uma região geográfica R e F_0, F_1, \dots, F_n conjuntos de geo-campos definidos sobre R , tais que o contradomínio das funções de todos os geo-campos em G_i é V_i , $i=0, \dots, n$.

⁸Estas operações podem envolver modificação da topologia (e.g. uma reclassificação é usualmente combinada com uma junção topológica).

Seja $\xi: V_1 \times \dots \times V_n \rightarrow V_0$. A operação pontual $\Xi: F_1 \times \dots \times F_n \rightarrow F_0$ induzida por ξ é tal que:

$$\forall f_1 \in F_1, \dots, f_n \in F_n,$$

$$\Xi(f_1, \dots, f_n) = f_0 \Leftrightarrow f_0(p) = \xi(\lambda_1(p), \dots, \lambda_n(p), \forall p \in R.$$

Dependendo dos domínios e contradomínio dos mapas de geo-campos, podemos considerar diferentes possibilidades para Ξ :

- *Operações unárias*: a entrada é um único geo-campo. Também são chamados *operações de transformação*, pois a operação equivale a um mapeamento entre os contradomínio dos campos de entrada e saída.
- *Operações booleanos*: são utilizados em análise espacial qualitativa e geram um TEMÁTICO, a partir de regras aplicadas a geo-campos (que podem ser instâncias de TEMÁTICO, NUMÉRICO ou DADO_SENSOR_REMOTO). As regras especificam o conjunto de condições a ser satisfeitas para cada tema de saída.
- *Operações matemáticas*: funções aritméticas, logarítmicas e trigonométricas, aplicadas a MNTs e a DADO_SENSOR_REMOTO. Podem gerar MNT, DADO_SENSOR_REMOTO ou TEMÁTICOS.

Dentro dos operadores matemáticos, vale destacar as seguintes subclasses:

- *operações de processamento de dados de sensoriamento remoto*: subclasse de operadores matemáticos, onde a entrada é um DADO_SENSOR_REMOTO e a saída é um DADO_SENSOR_REMOTO.
- *operações de classificação de dados de sensoriamento remoto*: subclasse importante dos operadores matemáticos, onde as entrada são instâncias da classe DADO_SENSOR_REMOTO e a saída é um TEMÁTICO.

A tabela 5.1 descreve os principais tipos de operações pontuais unárias (também chamados operações de transformação).

TABELA 5.1
OPERAÇÕES DE TRANSFORMAÇÃO

F_1 - entrada	F_2 - saída	Nome da Operação
TEMÁTICO	MNT	Ponderação
TEMÁTICO	TEMÁTICO	Reclassificação
DADO SEN. REMOTO	TEMÁTICO	Fatiamento
MNT	TEMÁTICO	Fatiamento de classes

Alguns exemplos de operações de transformação:

- “Reclassificar um mapa de vegetação com as classes {Floresta Ombrófila Densa, Floresta Ombrófila Aberta, Campinarana, Floresta Estacional} em um mapa com as classes {Floresta Densa, Floresta Aberta}.”
- “Obter um mapa hipsométrico a partir de um mapa de altimetria com o mapeamento { (0-300m) → Planície, (300-500m) → Planalto, (>500m) → Serras}”.

A figura 5.1 mostra um exemplo do operações de ponderação (conversão de um mapa de solos em um mapa de solos ponderado). Neste caso, $V_1 = \{ Le, Li, Ls, Aq \}$, $V_2=[0.0,1.0]$ and ξ é o conjunto de pares ordenados $\{(Le \rightarrow 0.60), (Li \rightarrow 0.20), (Ls \rightarrow 0.35), (Aq \rightarrow 0.10)\}$.

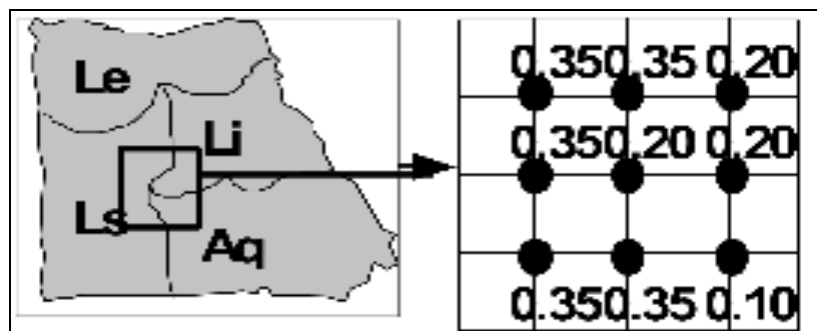


Figura 5.1 - Exemplo de operação de ponderação.

A figura 5.2 mostra um exemplo de uma operação de fatiamento em classes (conversão de um MNT em um TEMÁTICO) onde um mapa de declividade em graus é convertido para um mapa de classes de declividade a partir da transformação { (0-9) → “baixa”; (10-19) → “média”; (acima de 20) → “alta”}.

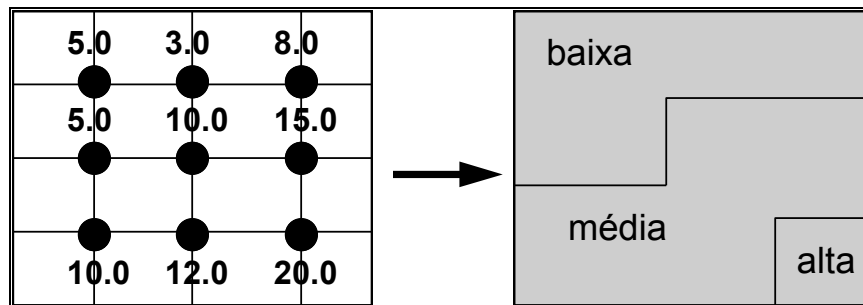


Figura 5.2 - Exemplo de operação de fatiamento em classes.

Como exemplo de *operação booleana*, tome-se o caso de determinar um mapa de aptidão agrícola, a partir dos mapas de solo, declividade, precipitação e do conjunto hipotético de regras expresso na tabela abaixo.

TABELA 5.2

REGRAS PARA APTIDÃO AGRÍCOLA

<i>Aptidão Agrícola</i>	<i>Solos</i>	<i>Precipitação média</i>	<i>Declividade</i>
Boa	Latosolo roxo ou Litossolo	> 100 mm	0-5
Média	Cambissolo	100-50 mm	5-10
Inapto	Aquoso	< 50 mm	>10

Como exemplo de *operação matemática*, tome-se a figura 5.3, onde f_1 é um mapa de solos ponderado e f_2 é um mapa de declividade (a declividade é o módulo das derivadas parciais da altimetria). A operação

$$\lambda_{\text{new}}(p) = \lambda_1(p) + 1/\lambda_2(p)$$

poderia ser utilizada como passo intermediário ao calcular um mapa de adequação de solos (quanto maior o valor, mais adequado).

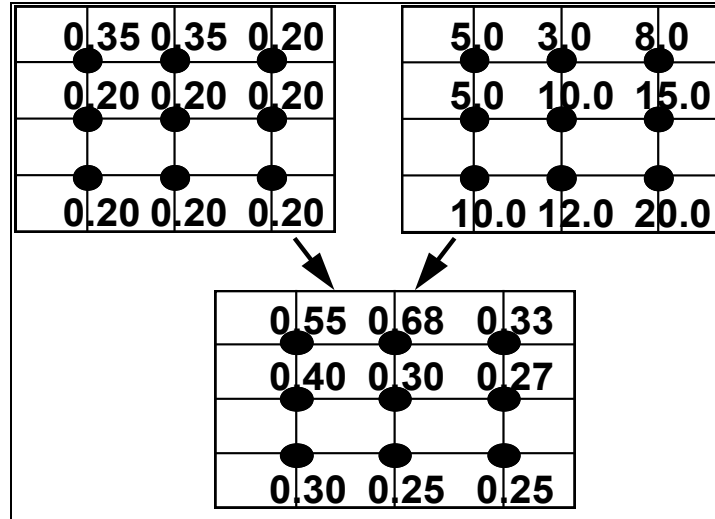


Figura 5.3 - Exemplo de uma operação matemática.

Operações de Vizinhança

Para definir as operações de vizinhança, é preciso introduzir os conceitos de região conectada, vizinhança e de tamanho de vizinhança.

Definição 5.2 Vizinhança em região geográfica.

Dada uma região geográfica R , um conjunto de pontos $P \in R$ é dito *conectado* se e somente se entre quaisquer dois pontos de P há uma linha ligando os dois pontos, inteiramente contida na região.

Dada uma região geográfica R , uma *vizinhança* é um mapeamento

$VL: R \rightarrow 2^R$, tal que $\forall p \in R, p \in VL(p)$ e $VL(p)$ é conectada.

Definição 5.3. Operações de Vizinhança em geo-campos.

Seja R uma região geográfica e sejam F_0 e F_1 conjuntos de geo-campos definidos sobre R e cujo contradomínio de suas funções é $V_i, i = 0, 1$.

Sejam $VL: R \rightarrow 2^R$ e $v: 2^{V_1} \rightarrow V_0$.

A operação de vizinhança $\Psi: F_1 \rightarrow F_0$ induzida por v é tal que

$\forall f_1 \in F_1, \Psi(f_1) = f_0 \Leftrightarrow f_0(p) = v(\{\lambda_1(x) \mid x \in L(p)\}), \forall p \in R$.

Nesta classe de operações, dado um geo-campo f_1 , o geo-campo de saída f_0 é computado baseado na dimensão e forma de uma vizinhança $L(p)$ em torno de cada localização p . Exemplos incluem:

- cálculos de valores *mínimo*, *máximo*, *médio*, *modal* para uma vizinhança em torno de um ponto.
- *filtros* para processamento de DADO SENSOR REMOTO.
- *métodos de interpolação espacial* para MNT (como médias por vizinho mais próximo).
- *mapas de declividade e exposição* para MNT.
- *índices de diversidade* para TEMÁTICO (onde o valor de saída está associada ao número de vizinhos de um ponto de entrada de uma classe que pertencem a classes distintas).

Como exemplo de operação de vizinhança, tomemos o caso do índice de diversidade, computado a partir de uma vizinhança 3 x 3 em torno de cada ponto. A idéia é que a biodiversidade de uma região é maior em áreas de contato ecológico entre regiões homogêneas. A figura 5.4 apresenta um mapa de vegetação e mostra o índice de diversidade computado para uma parte do mapa.

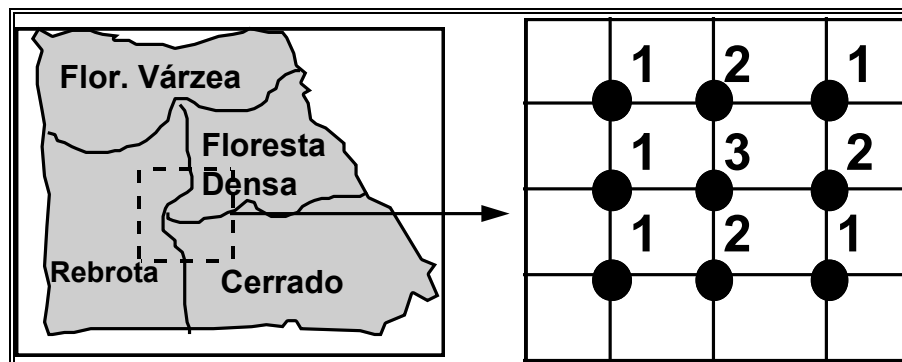


Figura 5.4 - Exemplo de operação “índice de diversidade”.

Operações Zonais

As operações zonais são sempre definidas sobre geo-campos das classes MNT ou DADO SENSOR REMOTO. Como a restrição desta operação pode ser um TEMÁTICO ou CADASTRAL, iremos considerar duas definições distintas. Nesta seção, apresentamos o caso de operações zonais onde a restrição é um TEMÁTICO. O segundo caso será discutido na seção 5.7.3 (Operações Mistas).

Inicialmente, precisamos definir o que entendemos por vizinhança zonal.

Definição 5.4 Vizinhança Zonal definida por um geo-campo temático.

Seja R uma região geográfica e F_t um conjunto de geo-campos definidos sobre R . Dado $f_t \in F_t$, a vizinhança zonal de f_t é um mapeamento $Z[f_t]: R \rightarrow 2^R$, tal que:

$Z[f_t(p)] = C_p \Leftrightarrow p \in C_p$ e $(\forall p \in C_p, f_t(x) = f_t(p))$ e C_p é conectado e C_p é o maior subconjunto de R com tais propriedades.

Em outras palavras, uma região zonal em torno de um ponto de um temático é o conjunto de pontos a ele conectado que possui o mesmo valor.

Definição 5.5. Operações zonais sobre geo-campos com restrição por mapa temático.

Seja R uma região geográfica e sejam F_1 e F_t conjuntos de geo-campos definidos sobre R , onde F_t é um geo-campo temático.

Seja $f_t \in F_t$ e $Z[f_t]: R \rightarrow 2^R$ a vizinhança zonal de f_t . Seja v uma função cujo domínio é 2^R . A operação zonal $\Gamma: F_1 \rightarrow F_1$ induzida por v é tal que,

$$\forall f_1 \in F_1, \Gamma(f_1) = f_2 \Leftrightarrow f_2(p) = v(\{\lambda_1(x) \mid x \in Z[f_t(p)]\}), \forall p \in R.$$

No caso de operadores zonais, um temático é utilizado para definir uma região de interesse aonde é computada a função. Os operadores zonais incluem:

- Média, máximo e mínimo e desvio padrão dos valores sobre uma região especificada.

- *Índice de variedade* dos valores, onde cada valor no mapa de saída será computado a partir do número de valores diferentes do geo-campo de entrada que pertencem a uma mesma região zonal.

Consideremos, por exemplo, a operação MÁXIMO ZONAL, onde tomamos um TEMÁTICO como restrição sobre um MNT, e retornamos o maior valor do MNT para cada tema (vide Figura 5.5).

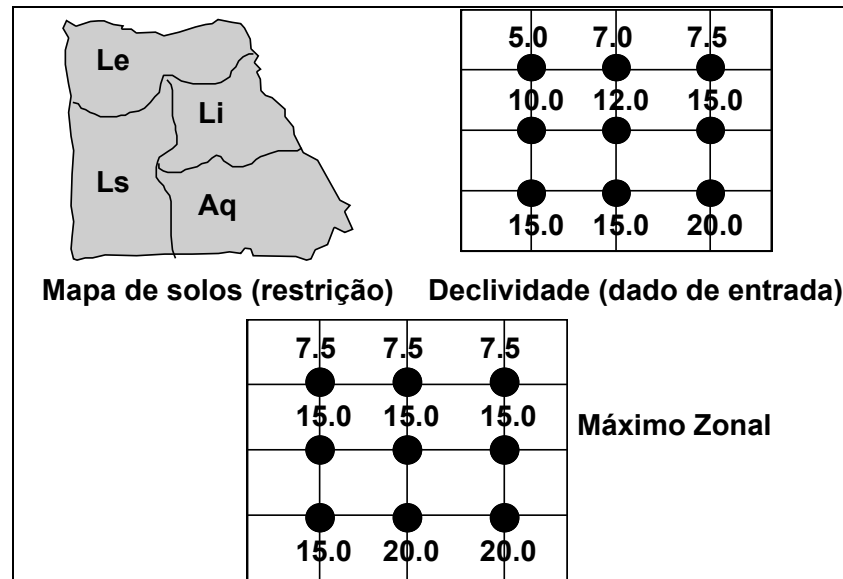


Figura 5.5 - Operação de máximo zonal.

Propriedades de geo-campos

As operações de propriedades sobre geo-campos podem ser divididas em três sub-grupos: propriedades pontuais, locais e zonais. A definição formal de cada uma destas é análoga às anteriores, com a diferença que a saída é um valor (ou conjunto de valores). Exemplos incluem cálculos estatísticos globais em um geo-campo: *média*, *mínimo*, *máximo*, *desvio padrão* e *histograma*.

5.5 RELACIONAMENTOS ESPACIAIS ENTRE GEO-OBJETOS

Em nosso modelo, os geo-objetos estão sempre associados a representações gráficas 2D (*pontos, linhas e regiões*). Como as operações da álgebra de geo-objetos podem envolver restrições espaciais, será fundamental caracterizar os relacionamentos espaciais, que podem ser divididos em (Güting, 1994):

- *relacionamentos topológicos*, tais como “dentro de” e “adjacente a”, invariantes a transformações biunívocas e bicontínuas (como as de escala, translação e rotação). Uma análise formal desta classe de relacionamentos foi proposta por Clementini et al. (1993), baseado no trabalho de Egenhofer e Herring (1990); tal proposta foi aprimorada pelo autor e está apresentada no que segue.
- *relacionamentos direcionais*, como “acima de” e “ao lado de”. Há uma grande variedade de propostas para este tipo de operadores, mas há pouca formalização neste campo.
- *relacionamentos métricos*, derivados das operações de *distância* e *direção*. O cálculo destas operações pressupõe sempre a existência de um espaço métrico, o que pode não ser sempre o caso.

Os relacionamentos direcionais representam um campo muito mais amplo para escolha de operadores do que os relacionamentos topológicos. Estamos diante de conceitos onde aspectos cognitivos (modelos de funcionamento da mente humana) devem ser levados em conta. Estudos recentes sobre o tema indicam que a formulação de conceitos espaciais pode variar de cultura para cultura (Frank and Mark, 1991).

A definição de um conjunto mínimo de operadores é objeto de muito debate na literatura:

- Freeman (1975) define um conjunto de 13 operadores: “à esquerda de”, “à direita de”, “acima” (mais alto que, sobre), “abaixo” (sob), “atrás”, “próximo a”, “longe de”, “ao lado de” (adjacente a), “tocando”, “dentro de”, “fora de”, “entre”.

- Feutchwangler (1989) lista seis: “adjacência”, “proximidade”, “subdivisão”, “sobreposição”, “vizinho mais próximo” e “sub-região”.
- Egenhofer e Herring (1987) usaram “disjunto”, “encontram”, “igual”, “dentro de”, “contém”, “cobre”, “coberto por” e “sobreposição”.
- Clementini et al. (1993) indicam cinco relacionamentos: “dentro de”, “superposto a”, “tocando”, “cruzando” e “disjunto”.

Dada a variedade de propostas, o esforço para formalizar os relacionamentos espaciais é parte fundamental da proposta de uma álgebra de geo-objetos. Para um conjunto bem-definido de objetos geográficos (casos “simples” de regiões sem buracos e linhas contínuas), a próxima seção apresenta uma formalização para os relacionamentos topológicos.

Formalização de Relacionamentos Topológicos

Apresenta-se nesta seção uma análise dos relacionamentos topológicos entre elementos do tipo ponto-linha-área, baseado nos trabalhos de Egenhofer (Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1992) e Clementini et al. (1993). Utilizamos os termos propostos por Clementini et al. (1993) que, a partir da análise das configurações possíveis entre os elementos ponto-linha-área, propõem cinco nomes para os relacionamentos topológicos: *touch*, *in*, *cross*, *overlap* e *disjoint*. No entanto, a metodologia usada para derivar as definições destes termos é mais rigorosa que a deste autor. No Apêndice 1, apresenta-se a prova de completeza destas definições, e cada caso possível de relacionamento é apresentado e classificado.

Consideramos que os elementos do espaço podem ser entendidos como *conjuntos de pontos* e os relacionamentos são formulados em termos de pontos, linhas e áreas “simples”:

- o espaço topológico é o R^2 ;
- um área é um conjunto de pontos 2D com um interior conectado, denotado por A° , uma fronteira conectada, denotada por δA , e um único

exterior conectado, denotado por \hat{A} . Assim, as áreas consideradas não tem “buracos”;

- as linhas são conjunto de pontos conectados, que podem ser ilhas (linhas circulares) ou possuem um ponto inicial e um ponto final distintos. A fronteira de uma linha (δL) é o conjunto vazio no caso de uma linha circular (ilha), e o conjunto dos pontos inicial e final nas demais situações. O interior (L°) de uma linha são os demais pontos.
- os elementos pontuais contém apenas um ponto. A fronteira δP de um ponto é sempre vazia.

Para caracterizar os relacionamentos entre dois elementos gráficos A e B do R^2 , (Egenhofer and Franzosa, 1991) propõem o uso de uma matriz 2 x 2, chamada *matriz de 4-intersecções*, que apresenta as relações topológicas binárias entre as fronteiras e o interior de A e as fronteiras e o interior de B:

$$\begin{bmatrix} \delta A \cap \delta B & \delta A \cap B^\circ \\ A^\circ \cap \delta B & A^\circ \cap B^\circ \end{bmatrix}$$

Considerando os valores “vazio” (\emptyset) e “não-vazio” ($\neg\emptyset$), para cada uma dos elementos da matriz, pode-se distinguir os relacionamentos entre os elementos gráficos. A matriz de 4-intersecções é capaz de modelar os relacionamentos entre duas áreas (sem “buracos”). Para o caso dos relacionamentos entre linhas e áreas, esta matriz é incapaz de diferenciar entre situações distintas. Por exemplo, os relacionamentos da figura 5.6 abaixo entre uma linha L e uma área A tem a mesma matriz de 4-intersecções, apesar de representar situações distintas (“inclusão” e “cruzamento”).

$$\begin{bmatrix} \delta L \cap \delta A & \delta L \cap A^\circ \\ L^\circ \cap \delta A & L^\circ \cap A^\circ \end{bmatrix} = \begin{bmatrix} \emptyset & \neg\emptyset \\ \neg\emptyset & \neg\emptyset \end{bmatrix}$$

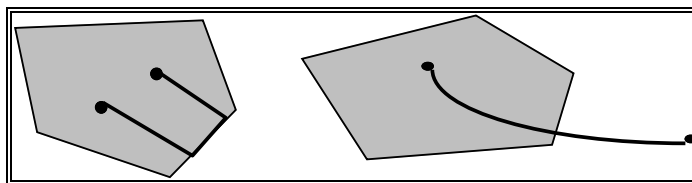


Figura 5.6 - Topologias linha-área com mesma matriz de 4-intersecções.

Em casos como este, será preciso também considerar o resultado da operação de intersecção entre os conjuntos de pontos que compõem a linha e a área. Se esta intersecção foi exatamente o conjunto de pontos que constitui a linha, então teremos o caso de “inclusão”. Caso contrário, ocorre a situação de “cruzamento”.

Igual consideração pode-se fazer para as relações “superposição” e “cruzamento” entre linhas, como ilustrado na figura 5.7. A matriz de 4-intersecções para os dois exemplos é

$$\begin{bmatrix} \delta L_1 \cap \delta L_2 & \delta L_1 \cap L_2^\circ \\ L_1^\circ \cap \delta L_2 & L_1^\circ \cap L_2^\circ \end{bmatrix} = \begin{bmatrix} \emptyset & \emptyset \\ \emptyset & \neg \emptyset \end{bmatrix}$$

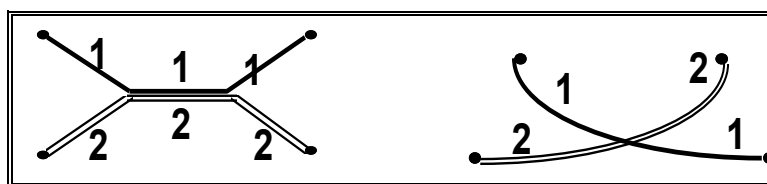


Figura 5.7 - Topologias linha-linha cujas matrizes de 4-intersecções são iguais.

Para distinguir entre os casos de “superposição” e “cruzamento”, será preciso lançar mão da idéia de dimensão da intersecção entre seus interiores, definida como se segue.

Definição 5.6 Dimensão de um conjunto de pontos

A *dimensão* de um conjunto de pontos ω é dada por:

$$\dim(\omega) = \emptyset \Leftrightarrow \lambda = \emptyset;$$

$$\dim(\omega) = 0 \Leftrightarrow \omega \text{ contém pelo menos um ponto e nenhuma linha ou área};$$

$$\dim(\omega) = 1 \Leftrightarrow \omega \text{ contém pelo menos uma linha e nenhuma área};$$

$$\dim(\omega) = 2 \Leftrightarrow \omega \text{ contém pelo menos um elemento do tipo área}.$$

No exemplo da figura 5.7, a dimensão da intersecção entre os interiores dos dois conjuntos - $\dim (L_1^\circ \cap L_2^\circ)$ - é igual a um (1) no caso de “superposição” entre as linhas e zero (0) no caso de “cruzamento”.

Relacionamento “toca”

Definição 5.7. Relacionamento “toca”

O relacionamento *toca* (aplicável a área/área, linha/área, linha/linha, ponto/área, ponto/linha) é tal que $\omega_1 \text{ toca } \omega_2 \Leftrightarrow (\omega_1 \cap \omega_2 \neq \emptyset) \wedge (\omega_1^\circ \cap \omega_2^\circ = \emptyset)$.

Diz-se que um conjunto de pontos ω_1 *toca* outro conjunto ω_2 se a única coisa em comum entre eles está contida na união de suas fronteiras, como ilustram os exemplos da figura 5.8.

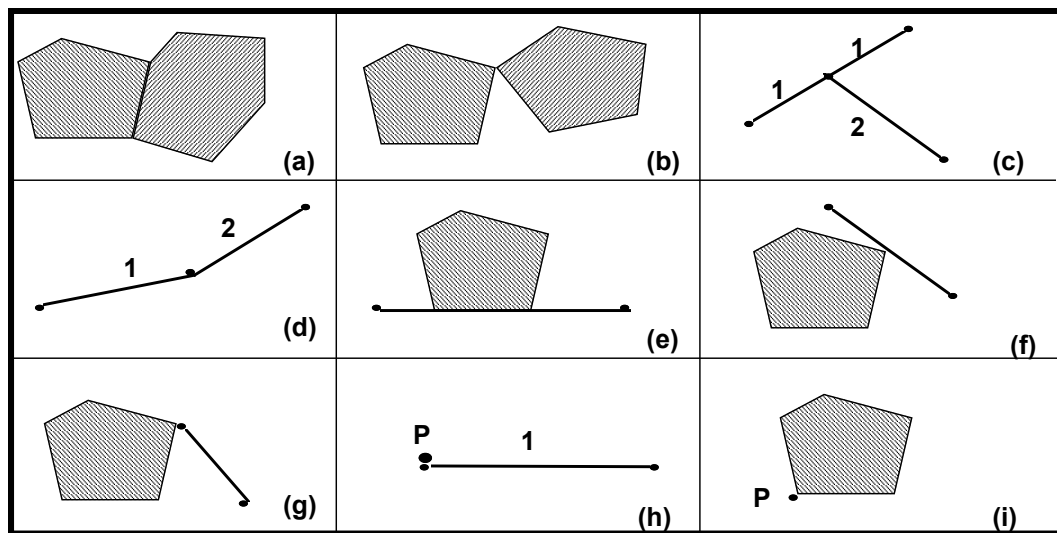


Figura 5.8 - Exemplos de situações topológicas que ilustram o relacionamento *toca*, no caso de duas áreas (a, b), duas linhas (c, d), linha e área (e, f, g), um ponto e uma linha (h) e um ponto e uma área (i). Adaptado de Clementini et al. (1993).

Relacionamento “dentro de”

Definição 5.8. Relacionamento “dentro de”

O relacionamento *dentro de* (aplicável a situações área/área, linha/área, ponto/área e ponto/linha) é tal que: $\omega_1 \text{ dentro de } \omega_2 \Leftrightarrow \omega_1 \cap \omega_2 = \omega_1$.

Diz-se que um conjunto de pontos ω_1 *está dentro de* outro conjunto ω_2 quando a intersecção dos dois conjuntos de pontos é o próprio ω_1 (vide figura 5.9).

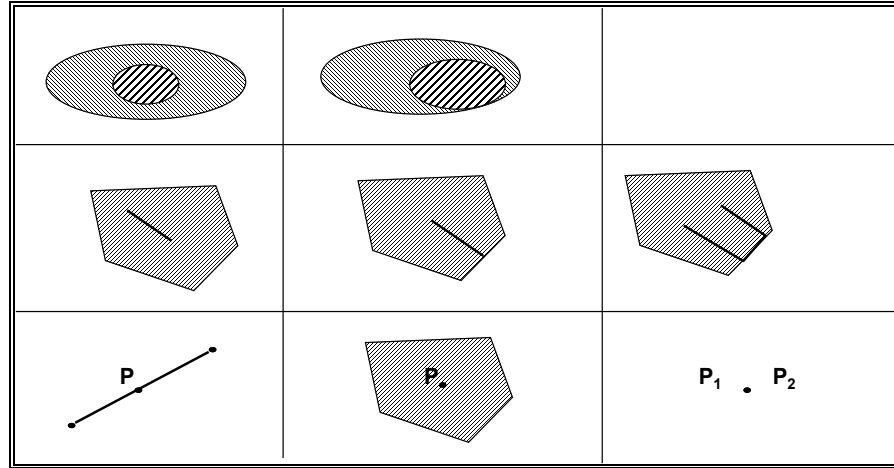


Figura 5.9 - Exemplos do relacionamento “dentro de” (contido em)

Relacionamento cruza

Definição 5.9 Relacionamento “cruza”

O relacionamento *cruza* (aplicável aos casos linha/área e linha/linha) é tal que:

$$L \text{ cruza } A \Leftrightarrow (L^\circ \cap A^\circ \neq \emptyset) \wedge ((L \cap A) \neq L)$$

$$L_1 \text{ cruza } L_2 \Leftrightarrow (L_1^\circ \cap L_2^\circ \neq \emptyset) \wedge (\dim(L_1 \cap L_2) = 0)$$

Duas linhas se *cruzam* se sua intersecção ocorre em ponto interno de ambas (note-se que a intersecção de seus pontos limites será definida como *touch*); de forma similar, uma linha cruza uma área se o interior da linha está parcialmente interno e parcialmente externo a esta área. Veja-se os exemplos da figura 5.10.

Relacionamento sobrepõe

Definição 5.10. Relacionamento “sobrepõe”

O relacionamento *sobrepõe*, aplicável a situações entre área/área, linha/linha e ponto/ponto, é tal que:

$$\omega_1 \text{ sobrepõe } \omega_2 \Leftrightarrow (\omega_1^\circ \cap \omega_2^\circ \neq \emptyset) \wedge (\dim(\omega_1^\circ \cap \omega_2^\circ) = \dim(\omega_1^\circ))$$

Diz-se que dois conjuntos de pontos ω_1 e ω_2 possuem *sobreposição* quando o resultado de sua intersecção é uma figura da mesma dimensão de ambos. Este relacionamento é aplicável a apenas a casos de elementos homogêneos. Veja-se os exemplos da figura 5.10.

Relacionamento *disjunto*

Definição 5.1.1 Relacionamento “disjunto”

O relacionamento *disjunto* vale em todas as situações e é tal que:

$$\omega_1 \text{ disjunto } \omega_2 \Leftrightarrow \omega_1 \cap \omega_2 = \emptyset$$

Duas figuras são disjuntas se sua intersecção é vazia, vide figura 5.10.

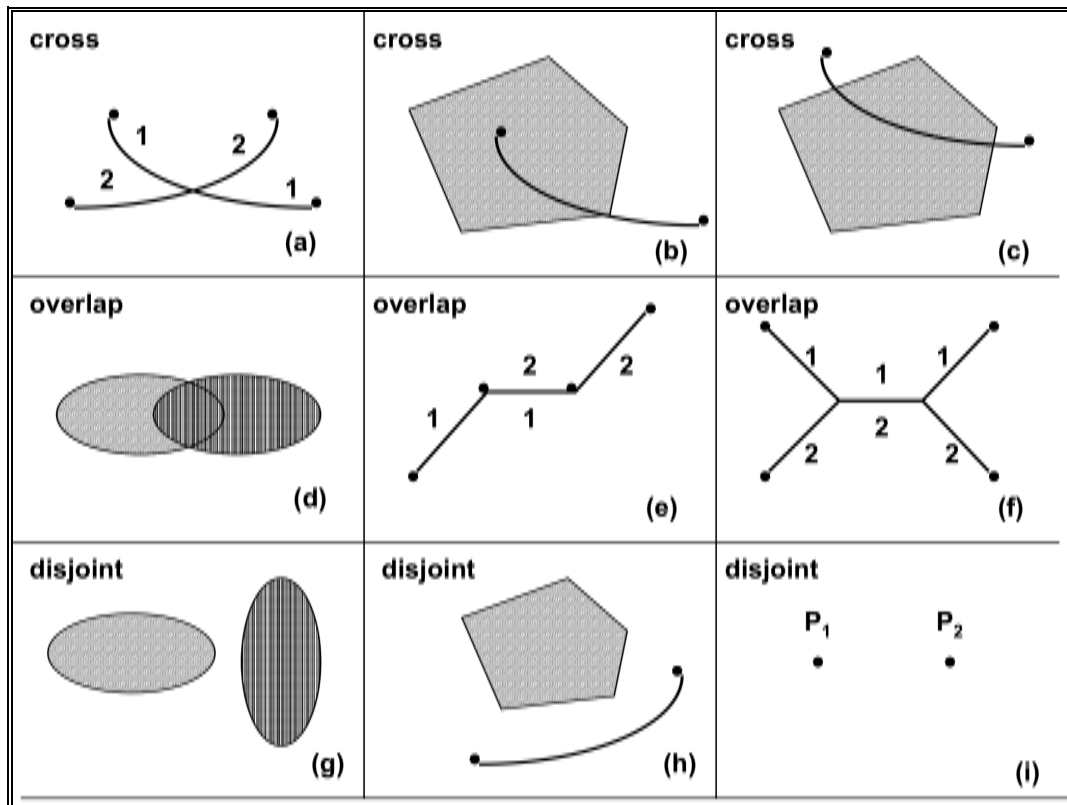


Figura 5.10 - Exemplos de relacionamentos: 1. *Cruza* entre duas linhas (a), linha e área (b, c). 2. *Sobreposição* entre duas áreas (d), duas linhas (e, f). 3. *Disjunto* entre duas áreas (g), linha e área (h), dois pontos (i). (Adaptado de Clementini et al. (1993))

5.6 OPERAÇÕES SOBRE GEO-OBJETOS

As operações sobre geo-objetos incluem:

1. *restrições sobre atributos*: computados em função dos atributos de entidades espaciais (e.g. “selecione todas as cidades de Alagoas com mortalidade infantil maior que 100  ”).
2. *restrições espaciais*: derivados a partir dos relacionamentos *topológicos* das entidades geográficas (e.g. “dê-me todas as escolas municipais do bairro Jardim Satélite”), de *direção* (“ao norte de”, “acima de”) ou *métricos* (e.g. “dê-me todas as escolas a menos de 500 m da Via Dutra”).
3. *propriedades de geo-objetos*: os resultados correspondem a predicados de um geo-objeto ou de um conjunto de geo-objetos (e.g. “calcule a média do valor venal das casas do bairro Jardim Esplanada” ou “indique o caminho ótimo para o ônibus que vai do Centro ao Jardim Uirá”).

Estas operações utilizam as primitivas definidas anteriormente: as relações topológicas *toca*, *dentro de*, *disjunto*, *cruza* e *sobrepõe*, as relações métricas unárias (comprimento, área, perímetro) e binárias (distância, direção).

Seleção por atributos

Definição 5.12 Seleção por atributos

O operador de seleção por atributos σ sobre um conjunto homogêneo GO , dada uma restrição ρ baseada apenas nos atributos descritivos de GO , é tal que:

$$\sigma_{\rho}(GO) = \{ go \in GO \mid \rho(go) \}.$$

Esta é operação semelhante à seleção da álgebra relacional, como indica o exemplo: "Recupere as cidades do Estado de São Paulo com população entre 100.000 e 500.000 habitantes".

Seleção espacial

Para definir as operações de consulta espacial, é necessário lançar mão do conceito de compatibilidade de predicado espacial.

Definição 5.13 Predicado espacial computável

Dados uma região geográfica R , um conjunto de geo-objetos GO e um objeto cadastral $oc = [R, GO, geo]$, um *predicado espacial* ξ sobre go_1 e go_2 *computável em* oc é uma restrição espacial, definida através de um relacionamento topológico (*inside*, *touch*, *cross*, *overlap* e *disjoint*) ou de um relacionamento métrico (*distância*), entre as representações $geo(go_1)$ e $geo(go_2)$ destes geo-objetos no objeto cadastral oc .

Por extensão, dada uma região geográfica R , dois conjuntos de geo-objetos GO_1 e GO_2 e dois objetos cadastrais $oc_1 = [R, GO_1, geo_1]$ e $oc_2 = [R, GO_2, geo_2]$, um *predicado espacial* ξ sobre $go_1 \in GO_1$ e $go_2 \in GO_2$ *computável em* (oc_1, oc_2) é uma restrição espacial, definida através de um relacionamento topológico ou métrico, entre as representações $geo_1(go_1)$ e $geo_2(go_2)$ destes geo-objetos, respectivamente, nos objetos cadastrais oc_1 e oc_2 .

Intuitivamente, os predicados espaciais que utilizaremos em nossas operações envolvendo geo-objetos são assertivas do tipo “rio que cruza o município de São José dos Campos, no mapa do Vale do Paraíba”.

Definição 5.14. Seleção espacial

Seja uma região geográfica R e um conjunto de geo-objetos GO . Seja ainda um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO numa região geográfica R . Dado um *predicado espacial* ξ computável em oc , o operador de seleção espacial $\varphi: GO \rightarrow GO$ é tal que:

$$\varphi_{\xi}(GO) = \{ go \in GO \mid \xi(geo(go)) \}.$$

O resultado desta operação é um subconjunto do conjunto original, composto de todos os geo-objetos que satisfazem o predicado geométrico, como ilustrado no exemplo:

- “selecione todas as regiões da França adjacentes à região de Midi-Pirenées (que contém a cidade de Toulouse)”.

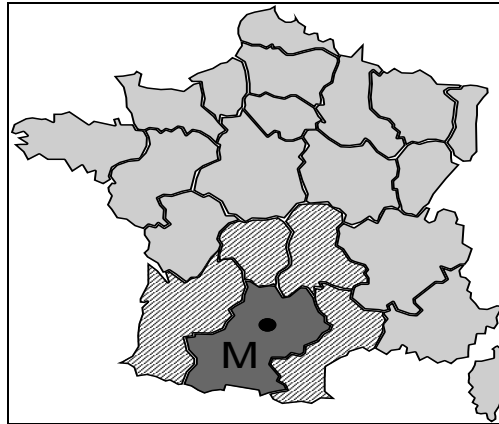


Figura 5.11 - Exemplo de operação de seleção espacial.

Junção espacial

Definição 5.15. Junção Espacial

Seja uma região geográfica R e dois conjuntos de geo-objetos GO_1 e GO_2 . Seja ainda dois objetos cadastrais oc_1 e oc_2 que mapeiam, respectivamente, objetos de GO_1 e GO_2 numa região geográfica R . Seja ainda um predicado espacial ξ computável em (oc_1, oc_2) .

A operação de junção espacial $\theta: GO_1 \times GO_2 \rightarrow GO_1 \times GO_2$ é tal que:

$$\theta_{\xi}(GO_1, GO_2) = \{ (go_1, go_2) \in (GO_1, GO_2) \mid \xi(geo_1(go_1), geo_2(go_2)) \}.$$

O termo *junção espacial* é empregado por analogia à operação de *junção* em banco de dados convencionais e denota o conjunto de operações onde ocorre a comparação entre dois conjuntos de objetos, baseado num predicado espacial computado sobre suas representações. Exemplos:

- “Para cada estrada da Amazônia, ache as reservas indígenas a menos de 5 km de uma estrada”.
- “Para as cidades do sertão cearense, ache quais estão a menos de 10 km de algum açude com capacidade de mais de 50.000 m³ de água”.

O resultado é uma *coleção de objetos e valores*, que satisfazem a restrição espacial. No primeiro exemplo, a resposta é um conjunto de pares (reserva, estrada) no segundo um conjunto de pares (cidade, açude).

Propriedades de Geo-Objetos

Neste caso, estão incluídos os operadores matemáticos tradicionais sobre os atributos dos geo-objetos (tais como média e desvio padrão) e também os operadores métricos (área e perímetro). Exemplos seriam:

- “Para as fazendas do Rio Grande do Sul, calcule a área média e o valor médio de ITR pago por cada uma”.
- “Liste os dez maiores consumidores de energia elétrica de Fortaleza”.

5.7 OPERAÇÕES ENTRE GEO-CAMPOS E GEO-OBJETOS

Analisaremos a seguir as operações que combinam geo-campos e geo-objetos, que apresentam particular interesse pois representam o vínculo entre as duas visões de dados em Geoprocessamento. Como os trabalhos da literatura abordam as operações geográficas privilegiando um dos pontos de vista, a ligação entre geo-campos e geo-objetos é tema ainda pouco explorado.

Geração de Geo-Objetos a partir de Geo-Campos

Definimos ainda duas grandes classes de operações: a operação de identificação e a operação de intersecção espacial.

Definição 5.16 Identificação.

Dada uma região geográfica R , seja F um conjunto de geo-campos temáticos definido sobre R , tal que o contradomínio das funções de todos os geo-campos em F é V .

Seja ainda um conjunto de geo-objetos $GO \subseteq A_1 \times \dots \times A_n \times 2^R$, cujo domínio do i -ésimo atributo $D(A_i)$ é V_i e cujos membros possuem localização em R , através de um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R .

A operação de identificação $\vartheta: G \rightarrow GO$ induzida por oc é definida por:

$$\forall f \in F, \vartheta_{oc}(f) = GO \Leftrightarrow \forall go \in GO, go = [a_1, \dots, v, \dots, a_n, geo(go)], \text{ e}$$

$$geo(go) = \{p \in R \mid f(p) = v\}.$$

Em outras palavras, a operação de identificação transforma um geo-campo temático em um objeto cadastral, que mapeia um conjunto de geo-objetos tais que um dos atributos de cada geo-objeto é o valor de geo-campo temático. O objeto cadastral terá a mesma representação geométrica do geo-campo temático que o originou.

Definição 5.17 Intersecção Espacial

Dada uma região geográfica R , sejam F_1, \dots, F_n conjuntos de geo-campos definidos sobre R , tais que o contradomínio das funções de todos os geo-campos em F_i é V_i , $i=1, \dots, m$. Seja ainda um conjunto de geo-objetos $GO \subseteq V_1 \times \dots \times V_m \times A_{m+1} \times \dots \times A_n \times 2^R$, cujos domínios de atributos $D(A_i)$ são V_i , $i=1, \dots, m$, e cujos membros possuem localização em R , através de um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R .

A operação de intersecção espacial $\otimes: F_1 \times \dots \times F_n \rightarrow GO$ induzida por oc é definida por:

$$\forall f_1 \in F_1, \dots, f_n \in F_n,$$

$$\otimes_{oc}(f_1, f_2, \dots, f_n) = GO \Leftrightarrow \forall go \in GO, go = [v_1, \dots, v_m, a_{m+1}, \dots, a_n, geo(go)], \text{ e}$$

$$geo(go) = \{p \in R \mid f_1(p) = v_1 \wedge \dots \wedge f_n(p) = v_n\}.$$

Esta classe de operações produz um *objeto cadastral* (e um conjunto de geo-objetos associados) a partir da *intersecção espacial* de um conjunto de geo-campos. Esta situação é típica de aplicações de *zoneamento*⁹, quando se faz a intersecção entre mapas temáticos para obter regiões homogêneas.

⁹Este exemplo (e toda a teoria de geração de mapas de geo-objetos a partir de geo-campos) foram inspirados pelo trabalho das equipes do IBGE, lideradas pela prof. Tereza Cardoso, que desenvolveu a metodologia do programa de Zoneamento econômico-ecológico.

Quando um objeto cadastral (e um conjunto de geo-objetos nele representado) é criado a partir da intersecção de geo-campos, cada geo-objeto resultante terá, como seus atributos descritivos, os valores de cada geo-campo de entrada (constante para cada geo-objeto). Veja-se o exemplo:

- “Determine as regiões homogêneas da Austrália, como cruzamento dos mapas de vegetação, geomorfologia e solos.” (Figura 5.12).

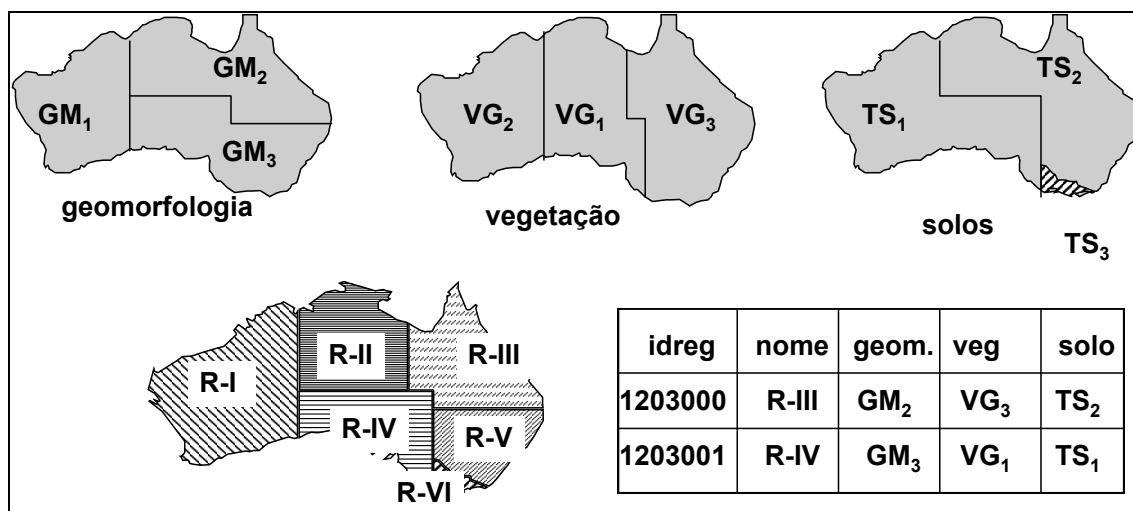


Figura 5.12 - Exemplo de intersecção espacial.

Nesta operação, pode ser conveniente permitir ao usuário que forneça um nome (“label”) que identifique a região. O atributo “nome” da tabela apresentada na figura seria então gerado pelo usuário, de forma individual para cada geo-objeto. Na literatura, a *intersecção espacial* é muitas vezes classificada erroneamente como “um tipo particular de junção espacial” (cf. Güting, 1994). Como vimos anteriormente, a operação de *junção espacial* parte de dois conjuntos de geo-objetos e produz, como resultado, pares de geo-objetos já existentes que satisfazem à restrição desejada. A intersecção espacial cria novos geo-objetos a partir de geo-campos. Deste modo, embora haja semelhanças entre os algoritmos gráficos utilizados para implementá-las, a operação de intersecção espacial (“*overlay*”) é conceitualmente diferente dos casos de *operações booleanas* entre geo-campos e *operações de junção espacial* entre geo-objetos¹⁰.

¹⁰Muitos sistemas comerciais (orientados para as estruturas gráficas) utilizam a mesma função (“*overlay*”) para implementar as três operações.

Geração de Geo-Campos a partir de Geo-Objetos

A partir de atributos (descritivos ou espaciais) de conjuntos de geo-objetos, pode-se gerar geo-campos. O novo mapa representa *uma restrição espacial definida a partir de um geo-objeto* ou *a variação de um atributo do conjunto de geo-objetos*, como ilustram os exemplos:

- “Gere um mapa das distâncias a partir da via Dutra na região de São José dos Campos.” (operação de *mapas de distância*).
- “para este conjunto de lotes, calcule um temático a partir do valor venal do terreno com as classes: temas A (até R\$ 300), B (de R\$ 300 a 1.000), C (de R\$ 1.000 a R\$ 4.000) e D (mais de R\$ 4.000).” (operação de *reclassificação por atributos*).

Definição 5.18 Mapas de distância.

Seja uma região geográfica R e F um conjunto de geo-campos definido sobre R tal o contradomínio de suas funções é \mathcal{M}^+ . Seja ainda um conjunto de geo-objetos GO , e um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R . A operação de mapas de distância $\Delta: GO \rightarrow G$ induzida por oc é tal que, dado o predicado espacial métrico $dist$ computável em oc e um objeto $go \in GO$:

$$\Delta_{oc}(go) = f \Leftrightarrow \forall p \in R, f(p) = dist(p, geo(go)).$$

Um *mapa de distâncias* é um mapa de geo-campos contendo as distâncias de cada ponto do mapa a um geo-objeto de referência (representado por um ponto, linha ou região). Trata-se de operação puramente geométrica (espacial). A figura 5.13 ilustra esta operação.

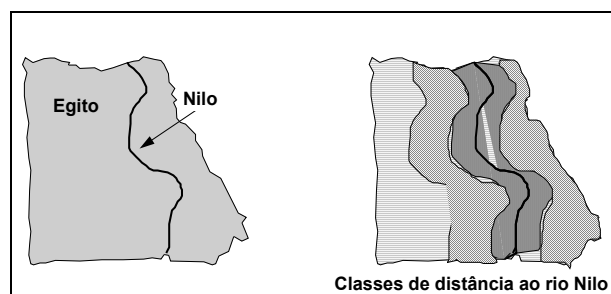


Figura 5.13 - Exemplo de mapa de distâncias.

Definição 5.19. Reclassificação por atributos

Seja uma região geográfica R , um conjunto de geo-objetos $GO \subseteq D(A_1) \times \dots \times D(A_n) \times 2^R$, e um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R .

Seja F um conjunto de geo-campos definido sobre R , tal que o contradomínio das funções de todos os geo-campos em F é $D(A_i)$, onde A_i é o i -ésimo atributo de GO .

A operação de reclassificação por atributos $\Omega: GO \rightarrow F$ induzida por oc é tal que:

$$\Omega_{oc}(GO) = f_0 \Leftrightarrow (\forall go \in GO, f_0(geo(go)) = go.A_i)$$

Alguns autores chamam a esta operação de *fusão* (Güting, 1994). A partir dos valores de um atributo específico dos geo-objetos de um mapa, obtém-se um geo-campo com a distribuição espacial deste atributo. Pode haver necessidade de recalcular a topologia do mapa resultante pois algumas regiões serão combinadas. Veja-se o exemplo ilustrado na figura 5.14:

“Para todos os países da América do Sul, gere um geo-campo temático com o crescimento demográfico de cada país, dividido em classes: $\{ (de\ 0\ a\ 2\ ao\ ano), (de\ 2\ a\ 3\ a.a.), (mais\ de\ 3\ a.a.) \}$.”

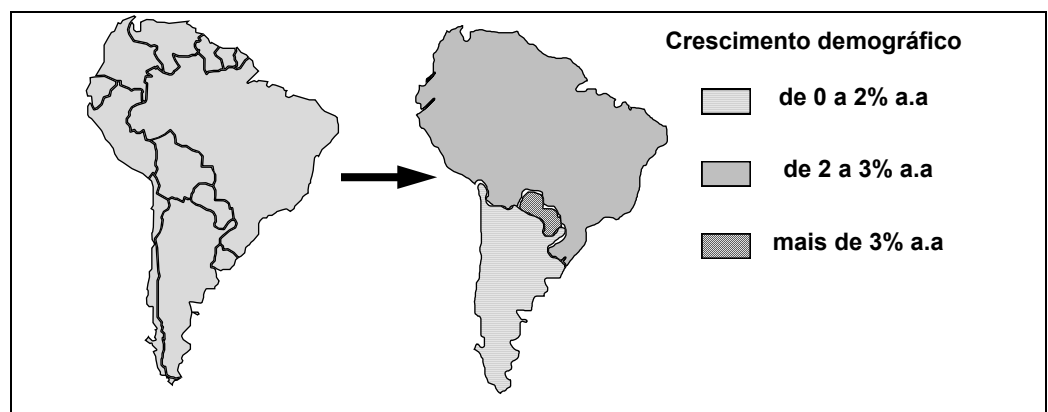


Figura 5.14 - Exemplo de *reclassificação por atributo*.

Operações mistas

Um conjunto importante de operações envolve operações sobre geo-campos aonde as restrições são dadas por geo-objetos (e vice-versa). No primeiro caso, pode-se pensar numa variante das *operações zonais* e no segundo, num outro tipo de junção espacial. Damos a seguir dois exemplos e a seguir a definição formal do primeiro caso (o segunda definição é análoga às anteriores).

1. *Operações zonais sobre geo-campos onde ge-objetos são restrições*: “Dados a altimetria e o mapa de municípios do Vale do Paraíba, crie um novo mapa aonde cada município será representado por sua altitude média”.
2. *Operações de seleção espacial aonde geo-campo é restrição*: “Dado um mapa de solos e um mapa de rios do Paraná, indique todos os rios que cruzam áreas com solos podzólicos”.

Inicialmente, definimos o que entendemos por vizinhança zonal obtida a partir de um objeto cadastral.

Definição 5.20 Vizinhança Zonal obtida a partir de um objeto cadastral.

Seja R uma região geográfica, um conjunto de geo-objetos $GO \subseteq D(A_1) \times \dots \times D(A_n) \times 2^R$, e um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R .

A vizinhança zonal de oc é um mapeamento $Z_{oc} : R \rightarrow 2^R$, tal que:

$\forall p \in R, Z_{oc}(p) = C_p \Leftrightarrow p \in C_p, (\exists go \in GO \mid C_p = geo(go)), C_p$ é conectado e C_p é o maior subconjunto de R com tais propriedades.

Em outras palavras, uma região zonal induzida por um geo-objeto é o conjunto de pontos de um temático é o conjunto de pontos a ele conectado que possui o mesmo valor.

A noção de vizinhança zonal obtida a partir de um objeto cadastral nos permitirá gerar a definição seguinte.

Definição 5.21. Operações zonais sobre geo-campos onde a restrição é um geo-objeto.

Seja uma região geográfica R , e F um conjunto de geo-campos definido sobre R , tal que o contradomínio das funções de todos os geo-campos em F é V . Seja um conjunto de geo-objetos $GO \subseteq D(A_1) \times \dots \times D(A_n) \times 2^R$, e um objeto cadastral $oc = [R, GO, geo]$, que mapeia objetos de GO em R . Seja $v : V \rightarrow V$. A operação zonal $\Gamma_o : F \times GO \rightarrow F$ induzida por oc e v é tal que :

$$\forall f_1 \in F, \Gamma_o(f_1, GO) = f_2 \Leftrightarrow f_2(p) = v \{Z_{oc}(p)\}.$$

5.8 RESUMO DAS OPERAÇÕES SOBRE GEO-CAMPOS E GEO-OBJETOS

Apresentamos a seguir um resumo das operações propostas, aplicáveis a geo-campos e geo-objetos, na Tabela 5.3. Estão indicados, para cada operação: a classe dos objetos de entrada e de saída, e dos objetos modificadores (quando cabível). Indicamos ainda as restrições de cada operação.

TABELA 5.3 RESUMO DAS OPERAÇÕES

Operação	Objeto Entrada	Objeto Modificador	Objeto Saída	Restrição
Ponderação	TEMÁTICO		NUMÉRICO	(função unária)
Fatiamento	NUMÉRICO		TEMÁTICO	(função unária)
Reclassificação	TEMÁTICO		TEMÁTICO	(função unária)
Booleana	NUMÉRICO, TEMÁTICO		TEMÁTICO	(regras)

Matemática	NUMÉRICO		NUMÉRICO	(fórmula)
Vizinhança	NUMÉRICO, TEMÁTICO		NUMÉRICO, TEMÁTICO	(função local e forma da vizinhança)
Zonais	NUMÉRICO	TEMÁTICO	NUMÉRICO	

Seleção Espacial	GEO-OBJETO (conjunto)	CADASTRAL	GEO-OBJETO (conjunto)	(predicado espacial)
Junção Espacial	GEO-OBJETO (conjuntos)	CADASTRAL	GEO-OBJETO e VALORES (conjunto)	(predicado espacial)
Identificação	TEMÁTICO		GEO-OBJETO (conjunto) CADASTRAL	
Intersecção Espacial	TEMÁTICO (n)		GEO-OBJETO (conjunto) CADASTRAL	
Mapa Distâncias	GEO-OBJETO	CADASTRAL	TEMÁTICO	(predicado métrico)
Reclassificaçã o Atributos	GEO-OBJETO (conjunto)	CADASTRAL	TEMÁTICO	(atributo)
Zonal sobre geo-objetos	TEMÁTICO, NUMÉRICO	GEO-OBJETO, CADASTRAL	TEMÁTICO, NUMÉRICO	
Seleção espacial (restr= geo- campo)	GEO-OBJETO (conjunto)	CADASTRAL, TEMÁTICO, NUMÉRICO	GEO-OBJETO (conjunto)	(predicado espacial)

6 CONSULTA, MANIPULAÇÃO E APRESENTAÇÃO DE DADOS GEOGRÁFICOS EM LEGAL

“Inventar uma linguagem poderia significar: inventar, com base em leis naturais (ou, em concordância com elas), uma aparelhagem para uma determinada finalidade; tem, porém., um outro sentido também, análogo aquele em que falamos na invenção de um jogo.”

L. Wittgenstein (Investigações Filosóficas).

6.1 APRESENTAÇÃO

O modelo de dados e as operações geográficas definidas anteriormente servem de base para a proposta de uma linguagem de consulta e manipulação espacial, denominada LEGAL (Linguagem Espacial para Geoprocessamento Algébrico). O objetivo de LEGAL é prover um ambiente geral para análise geográfica, incluindo operações sobre geo-campos e geo-objetos.

O que se fará neste capítulo é apresentar um esboço da sintaxe dos operadores desta linguagem e mostrar alguns exemplos de seu uso, deixando o detalhamento para trabalhos posteriores. Os operadores e palavras reservadas da linguagem estão apresentados em `COURIER MAIÚSCULAS`.

6.2 BREVE REVISÃO DA LITERATURA

Conforme observado anteriormente, a discussão das operações de manipulação (dita álgebra de mapas, cf. Tomlin, 1990) é feita, na literatura, de forma separada da análise das operações de consulta e apresentação (Egenhofer, 1989). Na maioria dos sistemas comerciais para Geoprocessamento, estas operações são implementadas em pacotes distintos.

Álgebra de Mapas de Tomlin

Para realizar operações de modelagem espacial em geo-campos, a abordagem mais corrente em sistemas de Geoprocessamento é utilizar a linguagem *MAP*

(Map Analysis Package), proposta por Dana Tomlin (1990). Esta proposta serve de base para muitos SIGs comerciais.

Em seu trabalho, Tomlin apresenta um conjunto de operações sobre mapas temáticos e modelos numéricos de terreno, definidos de forma a operar em geometrias matriciais, que incluem operações pontuais, de vizinhança (por ele chamadas de *focais*) e zonais.

A implementação mais difundida das idéias de Tomlin é a linguagem GRID, disponível no sistema ARC/INFO. Esta linguagem permite aos usuários realizar operações como¹¹:

```
OUTGRID = INGRID1 + INGRID2
OUTGRID = INGRID1 XOR 5
OUTGRID = SIN(INGRID1)*4/LOG(INGRID2)
```

A linguagem GRID, embora muito flexível, tem o sério inconveniente de não distinguir entre os diferentes tipos de operandos. Isto decorre do fato de ser fracamente tipada. Nos exemplos acima, se `INGRID1` for do tipo `MNT`, e `INGRID2` do tipo `TEMÁTICO`, o resultado pode não ter sentido. Em GRID, um mapa temático no formato matricial é chamado de “grade de inteiros”, e um modelo numérico de terreno de “grades de ponto flutuante”, o que estabelece uma confusão entre o dado geográfico e sua representação. Veja-se a seguinte explicação, retirada do manual do sistema ARC/INFO (ESRI, 1991):

“For both local and zonal functions, if any of the input grids are floating point, the output grid will contain floating-point values. If all of the input grids are integer grids, the output grid is an integer grid, with the exception of those local and zonal functions that are statistically based.”

Em função deste tipo de problemas, optamos por fazer de LEGAL uma linguagem fortemente tipada, estando os operadores ligados a um contexto

¹¹Utilizaremos o fonte tipográfico “courier” toda vez que nos referirmos a comandos a ser executados por um sistema computacional.

semântico, definido pelas diferentes especializações de geo-campos. Diferenças adicionais de LEGAL com relação às linguagens GRID e MAP incluem:

- disponibilidade de funções de tratamento de imagens;
- operações de classificação contínua e modelagem ambiental;
- chamada a funções externas, definidas pelo usuário.

Linguagens de Consulta Espacial

A maior parte das implementações de *linguagem de consulta* utiliza o padrão SQL para ter acesso aos atributos dos dados geográficos. No entanto, não é possível expressar relações topológicas diretamente em SQL, o que tem dado origem a muitas propostas de linguagens de consulta espacial (Egenhofer, 1989; Alves, 1990; OOi, 1990). Estas propostas são, de uma maneira geral, uma extensão da linguagem SQL com operadores que expressam relações espaciais. Cada linguagem tem seu conjunto próprio de operadores de extensão. As conclusões destas propostas de extensão da linguagem SQL são (Frank and Mark, 1991):

- estender SQL com operadores espaciais pode ser feito sem maiores problemas, desde que a semântica destes operadores tenha sido definida formalmente;
- as consultas em Geoprocessamento incluem restrições complexas, que requerem planejamento cuidadoso ao ser traduzidas para SQL;
- estender SQL para incluir controle de apresentação e apontamento é possível, mas ainda não existe qualquer sintaxe ou fluxo de ações compatível com a requerida pelos padrões cognitivos da linguagem natural.
- aconselha-se projetar uma linguagem separada para lidar com as questões de apresentação de dados, e não incluir estes comandos na cláusula SQL padrão "select-from-where".

A proposta do novo padrão para a linguagem SQL (SQL 3) contém um conjunto de extensões para lidar com dados espaciais. SQL 3 inclui tipos de dados para representar dados espaço-temporais, dados geométricos; seus operadores topológicos estão baseados nos trabalhos de Egenhofer e Franzosa (1991) e seus tipos de dados no padrão SAIF (Surveys and Resource Mapping Branch, 1993).

Para definir as operações de consulta espacial em LEGAL, partimos das idéias apresentadas por Egenhofer (1989), Egenhofer (1992), Egenhofer (1994), e Hemerly (1993a) e do padrão ODMG-93 (Cattell, 1994).

6.3 CARACTERÍSTICAS DE LEGAL

As principais características de LEGAL são:

- LEGAL é orientada-a-objetos; o resultado das operações é um conjunto de objetos ou um conjunto de valores.
- LEGAL é definida de forma semelhante ao padrão ODMG-93 (Cattell, 1994). Seguindo este padrão, as operações de consulta são implementadas utilizando uma sintaxe baseada em SQL.
- Tanto operações de consulta espacial quanto operações de manipulação fazem parte da sintaxe da linguagem.
- As operações sobre campos, as operações combinadas campos-objetos e a linguagem de apresentação são implementadas por operadores do mesmo nível semântico da linguagem SQL.

LEGAL é *fortemente tipada*. Cada dado pertence a uma das seguintes classes:

1. Classes básicas do modelo de dados, como segue:

- GEOOBJECT para instâncias de GEOOBJETO;
- NONSPATIAL para instâncias de NÃOESPACIAL;
- NETWORK para instâncias de REDE;
- CADASTRAL para instâncias de CADASTRAL;
- THEMATIC para instâncias de TEMÁTICO;
- IMAGE para instâncias de DADO SENSOR REMOTO;

- DTM para instâncias de MODELO NUMÉRICO DE TERRENO.

Estes dados tem os atributos básicos indicados na seção 4.5.5. No que segue, estes atributos básicos serão sempre expressos em maiúsculas.

2. Especializações das classes básicas, criadas através dos mecanismos de definição do esquema conceitual.

3. Classes auxiliares:

- PROJECT como partição geográfica do banco de dados, com os atributos básicos: PROJECTION (indica a projeção cartográfica), DATUM, LATLONG_BOX (indica as coordenadas geodésicas do retângulo envolvente), PROJ_BOX (indica as coordenadas de projeção do retângulo envolvente).

- TABLE para as tabelas de transformação entre geo-campos.

- MASK para máscaras aplicáveis a filtros.

- RECTANGLE e CIRCLE, utilizadas para definir vizinhanças.

4. Conjunto de objetos recuperados do banco:

- COLLECTION: agrupamento de objetos de classes diferentes.

5. Tipos de dados tradicionais INT, FLOAT, CHARACTER e BOOLEAN.

6. Tipo de dado MULTIMIDIA, usado para armazenar arquivos de texto, imagens, áudio e vídeo¹².

¹²Numa primeira aproximação, este tipo de dados poderia ser apenas um nome de arquivo contendo uma apresentação multimídia, que seria decodificado por seu conteúdo.

6.4 DEFINIÇÃO DE DADOS EM LEGAL

A parte de definição de dados em LEGAL inclui:

- criar e definir bancos de dados geográficos;
- definir o esquema do banco de dados;
- criar instâncias das classes deste esquema;

Definição de Esquema Conceitual

Em LEGAL, o esquema do banco de dados é criado pela especialização das classes básicas do modelo e das classes derivadas das classes básicas. A sintaxe

```
CREATE <especialização> (<atributos gerais>)
IS_A <classe básica> | <especialização>
      (atributos específicos)
```

será utilizada para especializar o modelo, como ilustrado nos exemplos seguintes:

- Criar a classe “Uso do Solo”, especialização de mapa temático, com os temas “Floresta Primária”, “Cerrado” e “Floresta Secundária”, e com os atributos “área” e “anos de regeneração”¹³.

```
CREATE Uso_do_Solo
( ATTRIBUTE nome      STRING;
  ATTRIBUTE ano       INT; }
IS_A THEMATIC
(THEMES = {"Floresta Primária",
"Cerrado", "Floresta Secundária"});
```

¹³O atributos indicados em maiúsculas indicam atributos herdados da classe básica.

- Criar a classe “Hospital”, especialização de `geo_objeto`, e com os atributos “nome”, “número de leitos”, “receita do SUS”.

```
CREATE Hospital
(ATTRIBUTE nome          STRING;
 ATTRIBUTE num_leitos    INT;
 ATTRIBUTE receita_SUS   FLOAT;)
IS_A GEOOBJECT;
```

- Criar a classe “Mapa Municipal”, especialização de `CADASTRAL`, com os atributos: número da planta, data do levantamento e empresa responsável.

```
CREATE Mapa_Municipal
(ATTRIBUTE nome          STRING;
 ATTRIBUTE num_planta    INT;
 ATTRIBUTE data_levantamento DATE;
 ATTRIBUTE empresa_respons STRING)
IS_A CADASTRAL;
```

A definição de classe pode admitir um valor por omissão (`DEFAULT`) que serve para inicializar as instâncias de forma conveniente. Vejamos um exemplo da criação de uma classe de imagens, que por omissão serão sempre instanciadas com resolução de 30m.

```
CREATE Landsat_TM
(ATTRIBUTE nome          STRING;
 ATTRIBUTE resolucao_h = 30 INT;
 ATTRIBUTE resolucao_v = 30 INT;
 ATTRIBUTE banda_espectral INT)
IS_A IMAGE;
```

Criação de geo-objetos e planos de informação

A criação explícita de instâncias de cada classe é obtida pelo comando `NEW`. Para o caso de geo-campos e de objetos cadastrais, será preciso ainda indicar qual a representação associada (matricial ou vetorial) e os parâmetros associados a cada representação. Tomemos os exemplos abaixo.

- Criar uma instância da classe `Uso_do_Solo` denotada por `acao_antropica_1990`, cuja representação matricial será uma grade de 100 x 100 metros de resolução, na escala 1:250.000.

```
VARIABLE acao_antropica_1990 Uso_do_Solo;

acao_antropica_1990 = NEW
(nome = "ação antrópica 1990",
 ano = 1990)
REPRESENTED_BY RASTER
(RESOLUCAO_H = 100 m,
 RESOLUCAO_V = 100 m,
 ESCALA = 1/250.000);
```

- Criar uma instância da classe `Hospital`, denominada `hospital_SJC`:

```
VARIABLE hospital_SJC Hospital;

hospital_SJC = NEW
(nome = "Hospital Municipal SJC",
 num_leitos = 300);
```

- Criar uma instância da classe `MAPA_MUNICIPAL`, para conter o mapa da cidade de São José dos Campos, cuja representação vetorial estará na escala 1:100.000, na projeção UTM e cujo retângulo envolvente possui coordenadas geográficas (-27°30", -45°00", -27°00", -44°30"),
-


```

VARIABLE mapa_SJC Mapa_Municipal;

mapa_SJC = NEW

(nome = "mapa SJCampos",

 num_planta = 25,

 data_levantamento = 1990,

 empresa_resp = "Aerodata")

REPRESENTED_BY VECTOR

(ESCALA = 100.000, PROJECAO = UTM,

 LAT_LONG_BOX= (-27°30", -45°00", -27°00", -44°30"));

```

6.5 CONSULTA ESPACIAL EM LEGAL

A consulta espacial em LEGAL tem dois componentes: uma *expressão de busca* expressa em SQL estendido e uma *resposta de consulta* que pode ser objeto de manipulação posterior.

A consulta espacial é aplicável a geo-objetos e geo-campos. No caso de geo-campos, dada a sua natureza contínua, os operadores disponíveis em LEGAL permitem apenas a recuperação baseada em seus atributos¹⁴. O mesmo procedimento vale para *geo-objetos complexos* (objetos cadastrais e de rede). Para os geo-objetos, LEGAL oferece mecanismos para recuperação baseada em atributos e em restrições espaciais.

Expressão de Busca

Em LEGAL, uma expressão de busca é construída a partir da linguagem OQL, com as seguintes extensões:

1. As restrições espaciais sobre geo-objetos serão sempre computadas a partir de suas representações geométricas num mapa de geo-objetos.

¹⁴Uma possível extensão incluiria aspectos de recuperação de geo-campos baseada em conteúdo, de forma análoga aos trabalhos recentes na área de "content-based image retrieval".

2. A cláusula `SELECT` pode permitir a recuperação de instâncias de classes em `LEGAL` e de valores de atributos destas instâncias.
3. A cláusula `FROM` contém a indicação das classes de objetos sobre as quais se realiza a consulta e do objeto cadastral (ou de rede) sobre o qual será computada a restrição espacial. Este objeto cadastral será indicado pelo qualificativo `ON MAP`, que pode ser omitido, no caso de haver apenas uma geometria presente para o geo-objeto.
4. A cláusula `WHERE` pode conter as restrições topológicas `INSIDE`, `TOUCH`, `CROSS`, `OVERLAP`, `DISJOINT` e a restrição métrica `DISTANCE`.

A formulação geral da expressão de busca em `LEGAL` será:

```
SELECT <valor > | <objeto> | <lista de objetos e
valores>

FROM <objeto> IN <classe> ON MAP <cadastral> | <rede>

WHERE <where_clause>
```

Tome-se o seguinte exemplo de consulta espacial em `LEGAL`:

- “Selecione todas as regiões da França a menos de 50 km da cidade de Toulouse”. Esta consulta está ilustrada na figura 6.1.

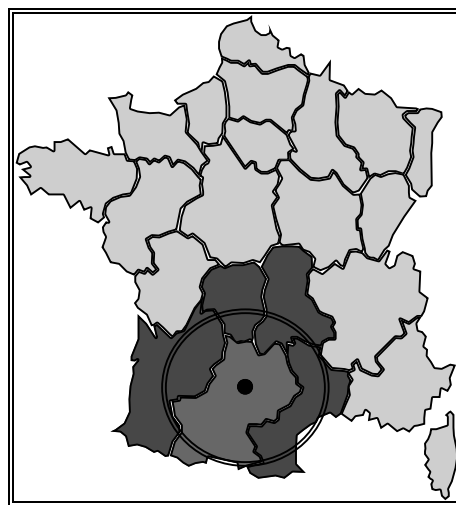


Figura 6.1 - Ilustração de consulta espacial.

A expressão de busca correspondente será:

```
SELECT regiao
FROM regiao IN Regioes_França ON MAP mapa_França,
      cidade IN Cidades_França ON MAP mapa_França
WHERE cidade.nome = "Toulouse" AND
      distance (regiao, cidade) < 50 km
```

Resposta de Consulta

Um aspecto que diferencia LEGAL da linguagem SQL tradicional é o procedimento para manipular os objetos resultantes de uma consulta. O resultado de uma consulta em SQL é uma tabela; para ter acesso aos itens resultantes de uma consulta, é utilizada a noção de “cursor”.

Este esquema de recuperação apresenta sérios inconvenientes para tratar dados espaciais. Em primeiro lugar, a manipulação posterior dos geo-objetos fica prejudicada e exige uma sintaxe não-intuitiva para a linguagem. A recuperação de respostas sob a forma de tuplas de valores também não permite tratar o caso de objetos complexos, como no caso de uma hierarquia de objetos.

Como visto no capítulo anterior, a resposta à consulta “selecione todas as reservas indígenas a menos de 30 km das represas da Eletronorte na Amazônia” é um objeto complexo composto dos pares de geo-objetos {(reserva, represa)}, passível de manipulação posterior.

Em resumo, *num bancos de dados geográficos é necessário recuperar geo-objetos e geo-campos, e não tabelas*. Para tanto, é preciso substituir o modelo de resposta de consultas implícito em SQL (“tabelas”) por um mecanismo orientado-a-objetos, em coerência com o modelo de dados proposto no capítulo anterior.

O modelo de resposta de consulta adotado em LEGAL é baseado nos conceitos do ODMG - Object Data Management Group (Catell, 1994). O modelo de dados do ODMG contempla a idéia de *coleções*, que contém um número

arbitrário de elementos de um mesmo tipo. A recuperação de um elemento dentro de uma coleção é feita por posição absoluta dentro da coleção ou a partir do elemento corrente.

Seguindo esta formulação, utilizaremos instâncias da classe `COLLECTION` para armazenar (de forma não-persistente) as respostas de uma consulta, com a sintaxe:

```
COLLECTION <colecão>
OF <classes_OBJETOS> | <classes_atributos>
```

Na linguagem `LEGAL`, a resposta a uma consulta pode devolver uma coleção de geo-objetos e de valores. Os valores correspondem a atributos de geo-objetos ou a propriedades destes objetos, computadas a partir de operadores espaciais como medidas de *distância*.

Para ilustrar a idéia, tome-se a consulta: "Recupere os postos de saúde num raio de 5 km do hospital municipal de S.J.Campos, indicando ainda sua distância ao hospital". Esta consulta será implementada como os resultado de duas consultas. No exemplo, supomos que as classes `Hospital` e `Postos_Saude` foram definidas previamente, como especializações de `OBJECT` e a classe `Cadastro_Urbano` é uma especialização de `CADASTRAL`.

```
COLLECTION postos_dist (Postos_Saude, FLOAT);
VARIABLE cad_SJC      Cadastro_Urbano;
cad_SJC = ( SELECT cad FROM cad IN Cadastro_Urbano
            WHERE cad.nome = "Mapa SJCampos");
postos_dist =
  (SELECT postos, distance (postos, hospital)
   FROM postos IN Postos_Saude ON MAP cad_SJC,
        hospital IN Hospital ON MAP cad_SJC,
        WHERE hospital.nome = "Municipal"
        AND distance (postos, hospital) < 5 km );
```

Exemplos de consultas espaciais

Apresentamos a seguir exemplos de consultas espaciais em LEGAL. Consideremos inicialmente a consulta: “Selecione todas as fazendas da família Magalhães cruzadas pela Linha Verde que liga Salvador a Aracaju, adjacentes a esta, ou a menos de 5 km dela”. Suporemos a existência das classes Fazenda, Estrada, Rede_Estradas, Cadastro_Fazendas.

```
VARIABLE estradas_BA          Rede_Estradas;
VARIABLE cad_faz_BA           Cadastro_Fazendas;
COLLECTION fazendas_Toninho   Fazendas;

// Recupere o mapa das fazendas
cad_faz_BA= (SELECT cad FROM cad IN Cadastro_Fazendas
              WHERE cad.nome = "mapa_faz_BA");

// Recupere as estradas da Bahia
estradas_BA= (SELECT rede FROM rede IN Rede_Estradas
              WHERE rede.nome = "mapa_estr_BA");

// Agora, as fazendas de Painho !!
fazendas_Toninho= (SELECT fazenda
                   FROM fazenda IN Fazendas ON MAP cad_faz_BA,
                   estrada IN Estradas ON MAP estradas_BA,
                   WHERE (estrada.name = "Linha_Verde" AND
                          fazenda.proprietario = "Magalhães" AND
                          (estrada TOUCH fazenda OR
                           estrada CROSS fazenda OR
                           distance(fazenda, estrada) <= 5km)));
```

O próximo exemplo ilustra o caso de uma ligação entre duas tabelas. Considere a consulta: “Selecione todas as fazendas cadastradas pelo INCRA na região do Pontal do Paranapanema, com mais de 1000 ha, e cujos proprietários pagam mais de 10.000 reais de ITR”.

Neste caso, a consulta deve combinar os objetos da classe FAZENDAS com o objeto não-espacial `cadastro` da classe `CADASTRO_INCRA`, onde está arrolado o cadastro alfanumérico de propriedades. A figura 6.2 ilustra esta situação, aonde os atributos dos geo-objetos e dos objetos não-espaciais foram implementados na forma de relações. A chave de ligação entre as duas relações é o número do cadastro INCRA (chave primária da segunda relação).

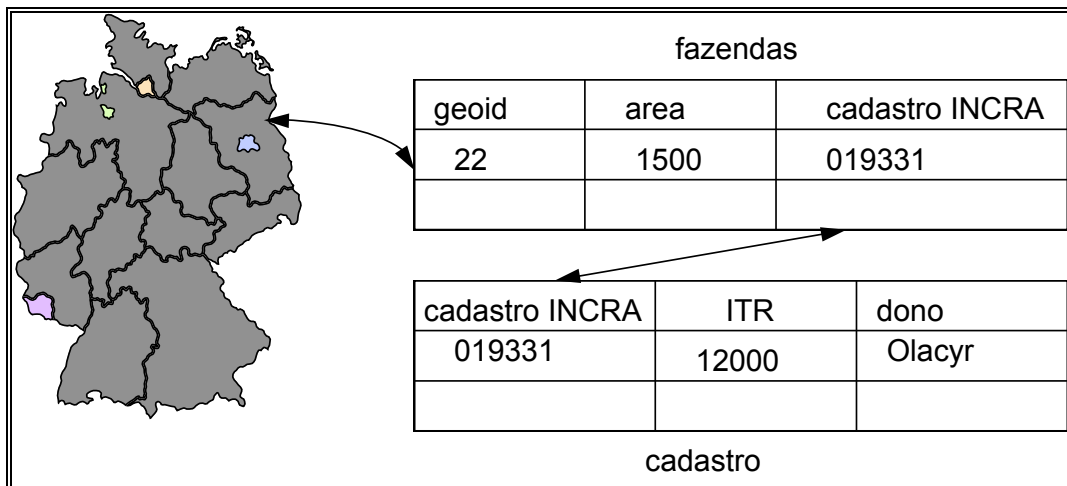


Figura 6.2 - Exemplo de consulta entre geo-objetos e objetos não-espaciais.

```

// Dados básicos existentes no Banco
VARIABLE mapa_faz_SP      Mapa_Fazendas;
VARIABLE mapa_reg_SP      Mapa_Regioes;
// Dados recuperados por selecao espacial
COLLECTION fazendas_PP    Fazenda;
// Recuperacao do mapa de fazendas
mapa_faz_SP= (SELECT mapa FROM mapa IN Mapa_Fazendas
              WHERE mapa.nome = "mapa_faz_SP");
// Recuperacao do mapa de regioes
mapa_reg_SP= (SELECT mapa FROM mapa IN  Mapa_Regioes
              WHERE mapa.nome = "mapa_regioes_SP");
// Recuperacao do objeto "regiao_pontal"
// Selecao das fazendas
fazendas_PP= ( SELECT fazenda
              FROM fazenda IN Fazenda ON MAP mapa_faz_SP,
              regiao IN Regioes ON MAP mapa_reg_SP,
              cadastro IN Cadastro_INCRA
              WHERE (fazenda.num = cadastro.num AND
                    cadastro.ITR >= 10.000 AND
                    fazenda.area >= 1.000 AND
                    regiao.nome = "Pontal do Paranapanema" AND
                    fazenda INSIDE regiao));

```

OPERAÇÕES DE JUNÇÃO ESPACIAL

A operação usual de junção espacial compara dois conjuntos de geo-objetos, computando um predicado espacial sobre suas representações em mapas de objetos. Sua sintaxe é análoga à junção convencional, como ilustram os exemplos a seguir. Consideramos também um caso especial de junção espacial, que ocorre quando comparamos um conjunto de geo-objetos com os valores de um geo-campo.

Junção Espacial sobre Conjuntos de Geo-Objetos

No primeiro caso, consideramos a consulta: “Mostre os aeroportos que estão dentro das reservas indígenas da Amazônia ou a menos de 50 km de uma reserva”. Suporemos a existências das classes `Aeroporto`, `Reserva`, `Mapa_Reserva` e `Mapa_Aeroportos`.

```
// definicao dos dados
VARIABLE mapa_reserv      Mapa_Reserva;
VARIABLE mapa_aerop      Mapa_Aeroportos;
// definições da coleção
COLLECTION aero_res (Aeroporto, Reserva, FLOAT);
// recuperacao dos mapas
mapa_reserv= (SELECT mapa FROM mapa IN Mapa_Reserva
              WHERE mapa.nome = "mapa_res_AMZ");
mapa_aerop=(SELECT mapa FROM mapa IN Mapa_Aeroporto
             WHERE mapa.nome = "mapa_aerop_AMZ");
// consulta sobre geo-objetos
aero_res= ( SELECT aeroporto,reserva,
            distance(aeroporto, reserva)
            FROM aeroporto IN AEROPORTO ON MAP mapa_aerop,
            reserva IN RESERVA ON MAP mapa_reservas
            WHERE (aeroporto INSIDE reserva
            OR distance (aeroporto, reserva) < 50 km));
```


.1 Junção Espacial entre Geo-Objetos e Geo-Campo

Neste caso, o predicado espacial a ser computado sobre um mapa de geo-objetos e um geo-campo, como na consultas indicadas a seguir. Para permitir o cômputo destas operações, definimos o operador *region*, que aplicado a um geo-campo, retorna o conjunto de pontos que possui um dado valor (temático ou numérico).

1. “Dado um mapa de solos e um mapa de rios do Paraná, indique todos os rios que cruzam áreas com solos podzólicos”

```
// definições dos dados
VARIABLE mapa_solos      Mapa_Solo;
VARIABLE rede_rios      Hidrografia;
// definição da coleção
COLLECTION rios_podz      Rios;
// recuperacao dos mapas
rede_rios= (SELECT rede FROM rede IN Hidrografia
            WHERE rede.nome = "mapa_rios_PR");
mapa_solos= (SELECT mapa FROM mapa IN Mapa_Solo
            WHERE mapa.nome = "mapa_solos_PR");
// consulta sobre geo-objetos
rios_podz= (SELECT rio
            FROM rio IN RIOS ON MAP rede_rios,
            WHERE rio CROSS region(mapa_solos= "podzólico"));
```

2. “Indique todos os lotes de Cubatão que estão em áreas com alta declividade”.

```
// definicao dos dados
VARIABLE campo_decl      Declividade;
VARIABLE mapa_lotes      Mapa_Lotes;
// definições da coleção
COLLECTION lotes_risco    Lotes;
// recuperacao dos mapas
campo_decl= (SELECT gc FROM gc IN Declividade
             WHERE gc.nome = "declivid_CBT");
mapa_lotes= (SELECT mapa FROM mapa IN Mapa_Lotes
             WHERE mapa.nome = "mapa_lotes_CBT");
// juncao espacial sobre geo-objetos e geo-campo
lotes_risco= ( SELECT lotes
              FROM lotes IN LOTES ON MAP mapa_lotes
              WHERE
                lotes INSIDE region (campo_decl = "Alta")
              OR
                lotes OVERLAP region (campo_decl = "Alta"));
```

Predicado espacial para redes

No caso de redes, motivados pela topologia arco-nó, é conveniente introduzir o predicado espacial `LINKED_TO`, que indica a ligação topológica entre dois objetos em uma rede. Trata-se de um predicado a ser utilizado na “where-clause” da operação de seleção espacial, computado a partir da sintaxe:

`<geo_objeto> LINKED_TO <geo_objeto>.`

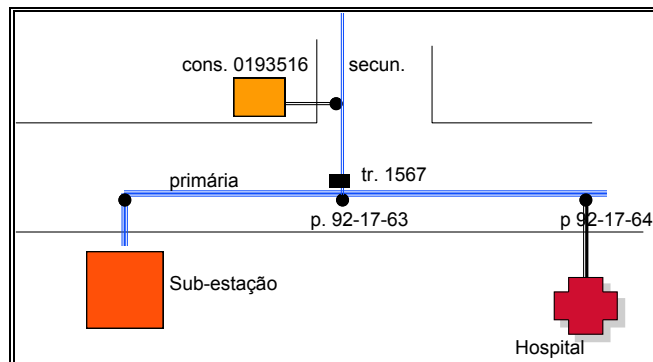


Figura 6.3 - Exemplo de rede elétrica.

Como exemplo, considere-se o caso ilustrado na figura 6.3, onde queremos saber “Quais os hospitais ligados à sub-estação do Jardim Botânico, que são da rede elétrica do bairro da Lagoa”. O programa em LEGAL seria:

```

COLLECTION cons      Consumidor;

VARIABLE rede_Rio  Rede_Eletrica;

rede_Rio= (SELECT rede FROM rede IN Rede_Eletrica
           WHERE rede.nome = "Rede_Rio");

hospitais= (SELECT cons
            FROM cons IN Consumidor ON MAP rede_Rio,
            sub_estacao IN Sub_Estacao ON MAP rede_Rio
            WHERE sub_estacao.nome = "Jardim Botânico"
            AND cons.type = "hospital"
            AND  cons LINKED_TO sub_estacao );

```

6.6 MANIPULAÇÃO ESPACIAL EM LEGAL

Após a recuperação dos elementos de um banco de dados geográficos, LEGAL permite a aplicação de operações de geo-campos e geo-objetos, definidas no capítulo 5, e que incluem:

- Transformações pontuais entre geo-campos.
- Operações booleanas entre geo-campos.
- Operadores matemáticos pontuais, tais como funções aritméticas e trigonométricas (aplicáveis a `IMAGE` e a `DTMs`).
- Métodos de classificação contínua, utilizando a função `FUZZY`, que permite gerar um mapa nebuloso (variando no intervalo $[0..1]$).
- Operações de vizinhança.
- Operações zonais entre geo-campos.
- Operações zonais entre geo-campos e geo-objetos.

Operadores de Transformação sobre Geo-Campos

Os operadores de transformação realizam o mapeamento entre as várias classes de geo-campos, a saber:

- `WEIGHT`: transforma um `THEMATIC` em um `DTM`;
- `SLICE`: transforma um `DTM` ou uma `IMAGE` em um `THEMATIC`;
- `RECLASSIFY`: transforma um `THEMATIC` em outro `THEMATIC` de classe distinta.

Como regra, estes operadores requerem que o usuário defina um mapeamento entre os geo-campos de entrada e de saída. Para tal fim, a linguagem permite ao usuário definir tabelas que descrevem os mapeamentos desejados, sob forma de uma classe `TABLE` que pode ser especializado em diferentes tipos (`WEIGHT_TABLE`, `SLICE_TABLE`, `RECLASSIFY_TABLE`).

Operação de Ponderação

Uma tabela do tipo ponderação possui dois atributos: `CLASS_IN` e `FUNCTION`. Um valor de `FUNCTION` é uma lista com um tema de entrada e um peso, denotados com a seguinte sintaxe:

`<tema_entrada> : <peso>`

Ao definir uma tabela de ponderação, não precisamos indicar a categoria de saída (que será sempre uma especialização de DTM) pois a operação poderá ser utilizada como passo intermediário de uma operação complexa, como mostremos posteriormente. O exemplo a seguir ilustra uma operação de ponderação, onde um mapa de solos dá origem a um mapa de solo ponderado.

```
VARIABLE uso_solo_CE          Uso_do_Solo;
VARIABLE tab_peso             WEIGHT_TABLE;
VARIABLE solo_peso_CE         Solo_Ponderado;
tab_peso= NEW
    (CLASSE_IN = USO_SOLO,
     FUNCTION = (Le: 0.2, Li : 0.4, Aq : 0.5));
solo_pond_CE = WEIGHT (uso_solo_CE, tab_peso);
```

Operação de Reclassificação

A operação de reclassificação envolve o mapeamento entre duas especializações da classe `THEMATIC` e é definida por uma tabela de reclassificação, que possui três atributos: `CLASS_IN`, `CLASS_OUT` e `FUNCTION`. A função de reclassificação é uma lista de temas de entrada com os correspondentes temas de saída, com a sintaxe:

`<tema_entrada>: <tema_saida>;`

Note-se que precisamos explicitar a classe dos dados de entrada e a de saída, para que, durante a execução, possamos checar a validade sintática da operação.

No exemplo a seguir, temos um mapa de cobertura do solo na Amazônia com diferentes classes {"Floresta Densa", "Floresta Várzea", "Rebrota", "Área Desmatada", "Cerrado"}. Este mapa temático será reclassificado para um novo mapa apenas com as classes {"Floresta", "Desmatamento", "Cerrado"}. Supomos a existência de duas classes: VEGETACAO e DESMATAMENTO.

```
VARIABLE cobertura_vegetal      Vegetacao;
VARIABLE desmatamento          Desmatamento;
VARIABLE tab_recl               RECLASSIFY_TABLE;
tab_recl= NEW (CLASS_IN = VEGETACAO,
               CLASS_OUT = DESMATAMENTO,
               FUNCTION = ("Floresta Densa" : "Floresta",
                           "Floresta Varzea": "Floresta",
                           "Rebrota" : "Desmatamento",
                           "Area Desmatada" : "Desmatamento",
                           "Cerrado" : "Cerrado"));
// recuperação dos dados
cob_vegetal= (SELECT gc FROM gc IN Vegetacao
              WHERE gc.nome = "mapa_Amapa");
desmatamento= NEW (nome = "desmat_Amapa")
                REPRESENTED_BY RASTER
                (RES_X = 100 m, RES_Y = 100 m);
desmatamento= RECLASSIFY (cobertura, tab_recl);
```

Operações booleanas sobre geo-campos

O comando `SWITCH` permite aplicar *operações booleanas* a todos os tipos de geo-campos. É necessário especificar um conjunto de condições a ser satisfeitas para cada classe de saída, conforme o exemplo abaixo, onde um mapa de Aptidão de Solos é calculado, baseado na média de chuvas, topografia e tipo de solo.

```
VARIABLE    campo_solos      Solo;
VARIABLE    campo_topo      Topografia;
VARIABLE    campo_chuva      Chuva;
CREATE Aptidao (ATTRIBUTE nome STRING)
  IS_A THEMATIC (THEMES = {"Boa", "Media", "Baixa"});
VARIABLE    campo_aptidao    Aptidao;
campo_topo= (SELECT gc FROM gc IN Topografia
              WHERE gc.nome = "Top92");
campo_chuva= (SELECT gc FROM gc IN Chuva
              WHERE gc.nome = "chuva92");
campo_solos= (SELECT gc FROM gc IN Solo
              WHERE gc.nome = "solos94");
campo_aptidao= NEW (nome = "Aptidao_94")
  REPRESENTED_BY RASTER(RES_X= 100 m, RES_Y= 100 m);
campo_aptidao= SWITCH
{ "Boa" :    mapa_solo.THEME = "Le"
              AND mapa_chuva >= 1000
              AND mapa_topo <= 1500;
  "Media" :  mapa_solo.THEME = "Aq"
              AND mapa_chuva >= 600
              AND mapa_topo <= 1000;
  "Baixa" :  OTHERWISE;
}
```

Operações matemáticas

As operações matemáticas sobre geo-campos que são especializações de DTM e IMAGE, incluem:

- operações aritméticas: soma (+), subtração (-), multiplicação (*) e divisão (/);
- funções matemáticas: seno (sin), cosseno (cos), tangente (tan), arco tangente (atan), logaritmo (log), exponencial (exp), raiz quadrada (sqrt);
- relações: menor que (<), maior que (>), menor ou igual (<=), maior ou igual (>=), igual (==), diferente (!=).

Como exemplo de operador matemático, considere a *equação universal de perda de solo*, que é utilizada em Ciências do Solo.

```
VARIABLE   precip_anual       Precipitacao;
VARIABLE   erodibilidade      Erosao;
VARIABLE   declividade        Altimetria;
VARIABLE   exposicao           Altimetria;
VARIABLE   cobertura_solo     Solo;
VARIABLE   indice_protecao     Indice;

perda_solo: = precip_anual x erodibilidade x
               declividade x exposicao x cobertura_solo x
               indice_protecao;
```


Operadores locais

Os operadores locais básicos incluem: média, desvio padrão, soma, máximo, mínimo e diversidade, calculados a partir de uma vizinhança em torno de cada ponto.

Para definir os operadores locais é necessário inicialmente definir o conceito de vizinhança. Uma *vizinhança* em LEGAL pode ser definida como um *retângulo* envolvendo cada ponto, onde devem ser dadas as dimensões horizontal e vertical ou como um *círculo* ao redor de cada ponto, aonde o raio (nas dimensões do geo-campo) deve ser fornecido. Assim, a sintaxe geral dos operadores locais é

`operador_local (geo-campo, LOCALREGION)`

onde a definição `LOCALREGION` pode ser implementada como um retângulo ou um círculo. No primeiro caso, teremos:

`operador_local (geo-campo, RETANGLE, <dim_h>, <dim_v>)`

e no segundo,

`operador_local (geo-campo, CIRCLE, <raio>)`

Os operadores locais incluem:

- `LOCAL_SUM (geo-campo, LOCALREGION)`: calcula a soma local dos pontos do geo-campo para uma vizinhança.
- `LOCAL_MEAN (geo-campo, LOCALREGION)`: calcula a média local dos pontos para uma vizinhança especificada.
- `LOCAL_MAX (geo-campo, LOCALREGION)`: determina o máximo local dos pontos do geo-campo, dada uma vizinhança.
- `LOCAL_MIN (geo-campo, LOCALREGION)`: determina o mínimo local dos pontos do geo-campo, dada uma vizinhança.
- `LOCAL_STDEV (geo-campo, LOCALREGION)`: calcula o desvio padrão dos pontos do geo-campo, dada uma vizinhança.

- **LOCAL_VARIETY (geo-campo, LOCALREGION):** determina o número de valores distintos do geo-campo, para uma vizinhança em torno de cada ponto.

Outra importante classe de operações inclui os *filtros*. Aplicáveis a geo-campos do tipo IMAGEM e MNT, os filtros computam uma equação local, calculada a partir de uma máscara em torno de cada ponto. Para tornar mais flexível o uso destas operações, definimos a existência de uma máscara, através da seguinte sintaxe:

```
VARIABLE    <nome da mascara>      MASK;
<nome_da_mascara>= NEW
    ( DIM_X = <dimensao_horizontal>,
      DIM_Y = <dimensao_vertical>,
      COEF = <lista de coeficientes>);
```

A lista dos pesos deve ser fornecida a partir do ponto superior esquerdo, da esquerda para a direita. No exemplo a seguir, definimos uma máscara para um filtro de realce de imagem: a imagem é subtraída de uma média local, aonde o ponto central tem um peso maior. Trata-se de um processo conhecido como “unsharp masking” com “addback”.

A partir da definição de uma máscara, podemos aplicar um filtro à uma IMAGEM (ou a um MNT) utilizando a sintaxe:

```
<geo-campo>= FILTER (geo_campo, MASK);
```

como mostra o exemplo a seguir.

```

VARIABLE      add_back      MASK;
VARIABLE      sjc_spotpan, sjc_realce Spot_PAN;
add_back= NEW (DIM_X = 3, DIM_Y = 3,
               COEF = (-1, -2, -1, -2, 20, -2,
                       -1, -2, -1));
sjc_spotpan= (SELECT imagem FROM imagem IN Spot_PAN
              WHERE imagem.nome = "sjc_spotpan");
sjc_realce= NEW (nome= "sjc_realce")
              REPRESENTED_BY RASTER
              (RES_X = 10, RES_Y = 10);
sjc_realce= FILTER(sjc_spotpan, add_back);

```

Outros operadores importantes incluem:

- **REFINE**: obter um DTM de maior densidade do que o existente, com diferentes técnicas de interpolação (linear, superfícies quadráticas ou de quinto grau). Sua sintaxe é:

```
<mnt saida>= REFINE (<mnt_entrada>, <processo>);
```

```
<processo>= LINEAR | BICUBIC | QUINTIC
```

- **SLOPE, ASPECT** : calculam as derivadas locais de uma superfície e obtêm, como resultado, o módulo (declividade) e a orientação (exposição de vertentes), com a sintaxe:

```
<mnt saida>= SLOPE (<mnt entrada>);
```

```
<mnt saida>= ASPECT (<mnt entrada>);
```

OPERAÇÕES ZONAIS

No caso de operadores zonais, precisamos indicar o campo aonde será calculado o operador e a restrição de cálculo (que pode ser um mapa temático ou um geo-objeto).

Os operadores zonais tem a seguinte sintaxe geral:

`<campo_zonal>= op_zonal(<geo_campo>, <restricao>)`

aonde

`<restricao>= <geo_campo> | <geo_objeto>`

As operações zonais incluem:

- `ZONAL_MAX`: determina o valor máximo do campo de entrada para cada região de restrição.
- `ZONAL_MIN`: determina o valor mínimo do campo de entrada para cada região de restrição.
- `ZONAL_MEAN`: determina o valor médio do campo de entrada para cada região de restrição.

Outro operador importante é:

- `ZONAL_STATS`: produz uma estatística dos valores para cada região. A saída é um relatório , indicando para cada região zonal, os valores máximo, médio e mínimo.

Operações de classificação contínua

O uso de técnicas de classificação contínua busca utilizar as noções de conjuntos nebulosos (“fuzzy”), para substituir os processos tradicionais de geração de mapas.

Este aspecto é particularmente evidente no manuseio de mapas temáticos para análises de meio-ambiente. Uma fronteira arbitrária, definida precisamente por uma linha, entre dois tipos de solo, representa erradamente o que é, na realidade, uma variação contínua (Burrough, 1986). Quando realizamos operações de superposição entre mapas temáticos, o erro inerente à divisão arbitrária dos mapas em áreas estanques é propagado.

Estudos realizados pelo Centro Nacional de Pesquisa em Solos da EMBRAPA evidenciaram que os processos tradicionais de análise geográfica (discretizar variáveis para posteriormente combiná-las) engendram uma grande perda de sensibilidade no resultado final.

A alternativa é trabalhar sempre com dados em representação contínua, e utilizar análises quantitativas sobre mapas geográficos. Isto equivale, na prática, a trabalhar sempre com modelos numéricos de terreno para representar variáveis espaciais como solo, geomorfologia, vegetação.

Para uma discussão em maior detalhe sobre o problema, veja-se Druck e Braga (1995).

Em LEGAL, está disponível o operador de transformação de um geo-campo numérico num campo nebuloso (“fuzzy”), cujos valores variam entre $[0, \dots, 1]$. Uma aproximação de uma função de pertinência nebulosa é dada pelas equações quadráticas $\mu_L(x)$ e $\mu_U(x)$, como segue:

$$\begin{aligned} \mu_L(x) &= 1 && \text{se } x \geq \beta, \\ \mu_L(x) &= 1/[1 + \alpha(x - \beta)^2] && \text{se } x < \beta. \end{aligned}$$

$$\mu_U(x) = 1 \quad \text{se } x < \beta,$$

$$\mu_U(x) = 1/[1 + \alpha(x - \beta)^2] \quad \text{se } x \geq \beta.$$

Na equação $\mu_L(x)$, o parâmetro β indica o valor máximo, acima do qual a pertinência “fuzzy” é considerada total (isto é, igual a 1). Abaixo deste valor, a função tem uma forma quadrática, dependendo da variação do parâmetro α .

Conversamente, na equação $\mu_U(x)$, o parâmetro β indica o valor mínimo, abaixo do qual a pertinência “fuzzy” é considerada total (isto é, igual a 1). Acima deste valor, a função tem uma forma quadrática, dependendo da variação do parâmetro α . A figura 6.4 ilustra a função $\mu_L(x)$ para o caso $\alpha = 1$ e $\beta = 3$.

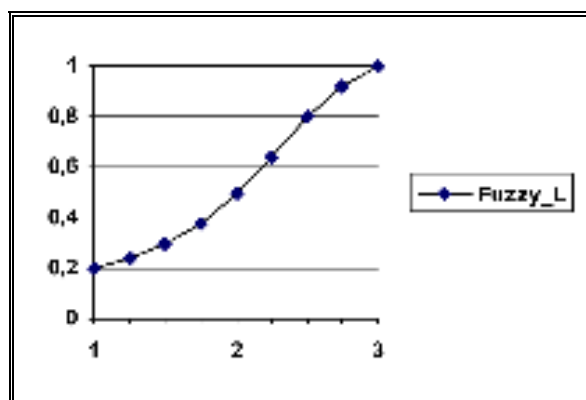


Figura 6.4 - Exemplo de função de pertinência “fuzzy”.

Em LEGAL, estão disponíveis duas operações que implementam as funções $\mu_L(x)$ e $\mu_U(x)$, com a seguinte sintaxe:

```
<mnt>= FUZZYL(<mnt_entrada>, <alfa>, <beta>);
```

```
<mnt>= FUZZYU(<mnt_entrada>, <alfa>, <beta>);
```

Em seu trabalho, Druck e Braga (1995) mostram um caso aonde se procura determinar classes de fertilidade de solos. As terras foram classificadas conforme sua exigência em termos de utilização dos insumos. A tabela 6.1 mostra os níveis de propriedades químicas para alguns parâmetros: cálcio e magnésio, fósforo e alumínio.

TABELA 6.1

<i>Propriedade</i>	<i>Classes Fertilidade</i>			
	1	2	3	4
Ca+++Mg++	Ca+Mg>3	2<Ca+Mg<=3	Ca+Mg<=2	Ca+Mg<2
P	P >= 30	10 < P < 30	P < 10	P < 10
Al++	Al > 0,3	0,3 < Al < 1,5	1,5 < Al < 4	Al > 4

Nesta tabela, a classe 1 indica o solo mais fértil e 4, o menos fértil. Para determinar as funções “fuzzy” correspondentes é escolhido o parâmetro β tal que a primeira classe de níveis de fertilidade de solo tenha o valor nebuloso 1; o parâmetro α é tal que o valor da função nebulosa $\mu_A(x)$ seja igual a 0,5 quando x tiver o valor inferior da segunda classe de fertilidade. A tabela 6.2 mostra os parâmetros “fuzzy” obtidos por este critério:

TABELA 6.2 - PARAMETROS FUZZY

<i>Propriedades</i>	α	β	Função
Ca+++Mg++	1	3	$\mu_L(x)$
P	0,0025	30	$\mu_L(x)$
Al++	0,3	0,694	$\mu_U(x)$

Cálculo de Propriedades

As operações de cálculo de propriedades objetivam obter estatísticas descritivas sobre os geo-campos e incluem:

- HISTOGRAMA: distribuição de frequência para as várias classes (ou valores) de um geo-campo e parâmetros estatísticos associados.
- TAB-CRUZ: tabulação cruzada é a distribuição de frequência mostrando as ocorrências comuns entre as classes de dois geo-campos.
- SPATIAL_AUTOC: calcula a autocorrelação espacial entre dois geo-campos.

6.7 TRANSFORMAÇÕES ENTRE GEO-CAMPOS E OBJETOS CADASTRAIS

Transformações de Campos para Objetos

Como discutido no capítulo anterior, uma das principais operações de transformação entre geo-campos e geo-objetos é a operação de indentificação., que transforma um geo-campo temático num objeto cadastral e cria geo-objetos representados neste mapa, onde um dos atributos de cada geo-objeto é o valor do geo-campo. Esta operação será realizada em dois passos. Inicialmente, cria-se um objeto cadastral que tem a mesma geometria do geo-campo temático, através da operação IDENTIFY_MAP, com a sintaxe:

```
<objeto_cadastral>= IDENTIFY_MAP (<temático>);
```

A seguir, será criado um conjunto de geo-objetos, no qual um dos atributos corresponde aos valores possíveis para o geo-campo temático, e que estará mapeado no objeto cadastral cuja geometria foi computada através da operação anterior. A operação que cria o conjunto de geo-objetos é IDENTIFY_OBJECTS, com a sintaxe:


```

<conjunto_geo-objetos>=
    IDENTIFY_OBJECTS(<temático>, <atributo>)
    ON MAP <cadastral>;

```

Nesta operação, o primeiro parâmetro indica o geo-campo temático a ser transformado. O segundo parâmetro indica o atributo dos geo-objetos que receberá o valor do tema do geo-campo temático (este atributo deverá ser do tipo “string”). A cláusula `ON MAP` denota o objeto cadastral criado anteriormente, cuja geometria (vetorial) é a mesma do geo-campo temático. A operação `IDENTIFY_OBJECTS` tem ainda a propriedade de tornar persistente uma coleção de geo-objetos.

Para ilustrar esta operação, consideremos a situação ao.nde deseja-se transformar um geo-campo temático que contém informação de solos, num conjunto de geo-objetos (chamadas aqui de *unidades de mapeamento*)¹⁵, mapeadas num objeto cadastral (*mapa de unidades*). O programa em LEGAL é mostrado, de forma simplificada, a seguir.

Inicialmente, temos as seguintes declarações de classes.

```

CREATE Campo_Solo (ATTRIBUTE nome STRING)
    IS_A THEMATIC (THEMES = "Lr", "Aq", "Cb",....);
CREATE unid_mapeamento
    (ATTRIBUTE nome          STRING,
    ATTRIBUTE tipo_solo      STRING,
    ATTRIBUTE ph             FLOAT)
    IS_A GEO-OBJECT;
CREATE mapa_unidades (ATTRIBUTE nome  STRING)
    IS_A CADASTRAL;

```

¹⁵Esta terminologia é empregada pelo Centro Nacional de Pesquisa em Solos da EMBRAPA.

O programa em LEGAL que realiza a operação tem a forma abaixo.

```
VARIABLE    solo_MG          Campo_Solo;

VARIABLE    mapa_unid_MG      Mapa_Unidades;

COLLECTION  unid_solo          Unidades_Mapeamento;

solo_MG = (SELECT gc FROM gc IN Campo_Solo
           WHERE gc.nome = "Solos_MG");

mapa_unid_MG = NEW (nome = "Mapa Unidades MG")
                REPRESENTED_BY VECTOR (SCALE = 50000);

mapa_unid_MG = IDENTIFY_MAP (solo_MG);

unid_solo = IDENTIFY_OBJECTS (solo_MG, tipo_solo)
           ON MAP mapa_unid_MG;
```

Um segundo tipo de operações de transformação de geo-campos em objetos cadastrais é da dado pela operação `INTERSECT`, que realiza a geração de mapas de objetos como regiões homogêneas obtidas por superposição de um mapa de geo-campos.

Esta operação deve ser vista como uma generalização do caso anterior. Também será realizada em dois passos. No primeiro passo, cria-se um objeto cadastral, cuja geometria corresponderá à representação vetorial resultante da intersecção espacial de um conjunto de geo-campos temáticos, através da operação `INTERSECT_MAP`, cuja sintaxe é:

```
<cadastral> = INTERSECT_MAP (<conjunto de geo_campos>);
```

A seguir, são criados os novos geo-objetos, mapeados no objeto cadastral gerado a partir da operação anterior. A sintaxe desta operação é:

```
<conjunto_geo-objetos>=
INTERSECT_OBJECTS (<geo-campos>, <atributo>,... )
ON MAP <cadastral>;
```

A operação `INTERSECT_OBJECTS` tem como parâmetros pares de geo-campos e atributos, o primeiro parâmetro de cada par indica o geo-campo temático a ser transformado. O segundo parâmetro indica o atributo dos geo-objetos que receberá o valor do tema do geo-campo temático (este atributo deverá ser do tipo “*string*”). A cláusula `ON MAP` denota o objeto cadastral criado anteriormente, cuja geometria (vetorial) é a mesma do geo-campo temático. A operação `INTERSECT_OBJECTS` tem ainda a propriedade de tornar persistente uma coleção de geo-objetos.

Como exemplo, tome-se um caso de zoneamento ecológico-econômico. Como parte deste tipo de estudo, uma fase inicial é obter, a partir dos condicionantes físicos, um mapa de regiões homogêneas (também chamadas *geofacies*) através da intersecção espacial dos mapas de geologia, solos e vegetação. Inicialmente, consideremos as seguintes definições de classes em `LEGAL`.

```
CREATE Campo_Vegetacao (ATTRIBUTE nome STRING)
    IS_A THEMATIC (THEMES = {"Floresta Ombrofila",
        "Mangue", "Campina",....} );
CREATE Campo_Geologia (ATTRIBUTE nome STRING)
    IS_A THEMATIC (THEMES = {"Formação Cambriana",
        "Metamorficas",....} );
CREATE Campo_Solo (ATTRIBUTE nome STRING)
    IS_A THEMATIC (THEMES = "Lr", "Aq", "Cb",....);
CREATE Regioes_Homogeneas
    (ATTRIBUTE nome          STRING;
    ATTRIBUTE tipo_solo      STRING;
    ATTRIBUTE tipo_veget     STRING;
    ATTRIBUTE tipo_geol STRING) IS_A GEO-OBJECT;
CREATE Mapa_Geofacies (ATTRIBUTE nome STRING)
    IS_A CADASTRAL;
```

A seguir, realizaremos as operações de intersecção.

```
VARIABLE geologia_AP      Campo_Geologia;
VARIABLE solos_AP        Campo_Solos;
VARIABLE vegetacao_AP     Campo_Vegetacao;
VARIABLE geofacies_AP    Mapa_Geofacies;
COLLECTION regioes_AP    Regioes_Homegeneas;
geofacies_AP= NEW (nome = "Mapa Geofacies Amapa",
  REPRESENTED_BY VECTOR
  (LATLONG BOX = [(02 00 00 S, 61 30 00 W),
                  (04 00 00 N, 60 00 00 W)]);
  SCALE = 1.000.000);
geo_facies_AP = INTERSECT_MAP (geologia_AP, solos_AP,
  vegetacao_AP);
regioes_AP =
  INTERSECT_OBJECTS (geologia_AP, tipo_geol,
  solos_AP, tipo_solo,
  vegetacao_AP, tipo_veget)          ON          MAP
geo_facies_AP;
```

Transformações de Objetos em Campos

As operações que nos permitem gerar geo-campos a partir de geo-objetos incluem:

- *Reclassificação por Atributos*: geração de geo-campo a partir de um atributo de um conjunto de geo-objetos. Sua sintaxe é:

```
<geo-campo>= RECLASS_ATR (<conjunto_geo-objetos>,
                             <atributo>) ON MAP <mapa_de_objetos>;
```

- *Distância*: geração de um mapa de distância a partir de um geo-objeto ou conjunto de geo-objetos, com sintaxe:

```
<geo-campo>=DISTANCE_MAP (<conjunto_geo-objetos>);
```

Para ilustrar estas operações, consideremos o caso onde um mapa de distribuição da pobreza do Brasil será obtido a partir dos dados do IBGE, agregados por estado. Suporemos que um dos atributos é a *renda per capita*. Para isto é necessário uma composição de operações: primeiramente é gerado um modelo de terreno com a distribuição da variável *renda per capita*, que será em seguida fatiado.

```

COLLECTION estados                Estado_Brasil;

VARIABLE  mapa_estados            Mapa_IBGE;

VARIABLE  campo_renda             Campo_Renda;

VARIABLE  tab_fatia               SLICE_TABLE;

ATTRIBUTE renda_per_capita        FLOAT;

mapa_renda:= NEW (nome = "Mapa Riqueza Brasil")

    REPRESENTED_BY RASTER

    (RES_X = 1000m, RES_Y = 1000m);

tab_fatia:= NEW ( CLASS_OUT = RENDA,

    FUNCTION =

    ([0-500]           : "Miseravel",

    (500-1000]        : "Pobre",

    (1000-2000]       : "Remediado",

    >2000             : "Rico"));

mapa_renda:= SLICE

    (RECLASS_ATR (estados, renda_per_capita)

        ON MAP mapa_estados), tab_fatia);

```

6.8 MECANISMOS DE FLUXO DE CONTROLE

Iteração sobre resposta de consulta

No modelo ODMG para coleções, adotado em LEGAL, é possível percorrer iterativamente os elementos de uma coleção. Para tal fim, é necessário definir um elemento do tipo `ITERATOR`, que mantém a posição corrente na coleção a ser percorrida, cuja estrutura é dada a seguir (Catell, 1994):

```

tipo ITERATOR <T>
propriedades:
    stable?: Boolean
    iteration_order: Enum (forward, backward)
operadores:
    next()-> element:T
    first()-> element:T
    last() -> element:T
    more?()-> element:T
    reset()
    delete()

```

Nesta estrutura, o operador `next` posiciona o iterador sobre o próximo elemento da coleção. Depois que `next` é aplicado ao último elemento da coleção, o valor do iterador será `nil`. A chamada inicial a `next` retorna o primeiro elemento da coleção, se o parâmetro `iteration_order` vale `forward`, e o último elemento, se o parâmetro `iteration_order` vale `backward`.

Em LEGAL, o tipo `ITERATOR` faz parte dos tipos básicos da linguagem. A associação entre um elemento do tipo `ITERATOR` e uma coleção é feita a partir da definição de uma instância desta classe:

```
ITERATOR <nome_iterador> OVER <coleção>;
```

O percorrimento sobre uma coleção será controlado por mecanismos do tipo `DO...UNTIL` e `WHILE`.

Para ilustrar este mecanismo, considere a operação “selecione todas as cidades do Ceará a menos de 50 km de açudes com capacidade de mais de 50.000 m³ de água”.

```
// Definicoes dos elementos do Banco
VARIABLE mapa_mun_CE      Mapa_Municipal;
INSTANCE mapa_acudes_CE   Mapa_Acudes;
// definições da coleção e do iterador
COLLECTION cid_acu_CE     (Cidade, Acude);
ITERATOR      elem  OVER      cid_acu_CE;
// recuperacao dos mapas
mapa_mun_CE= (SELECT map FROM map IN Mapa_Municipal
              WHERE map.nome = "mapa_Ceara");
mapa_acude_CE= (SELECT map FROM map IN Mapa_Acudes
               WHERE map.nome = "mapa_Acudes");
// consulta sobre geo-objetos
cid_acu_CE = (SELECT cidade, acude
              FROM cidade IN Cidade ON MAP mapa_mun_CE
               acude IN  Acude ON MAP mapa_acudes_CE
              WHERE acude.capacidade > 50.000
               AND DISTANCE (cidade, acude) < 50 km);
// percorre a coleção
WHILE MORE (elem)
BEGIN
    //processamento da colecao (nao mostrado)
    elem= NEXT (elem);
END
```


CONTROLE DE LAÇOS e TESTE DE CONDIÇÕES

As operações de manipulação requerem que se disponha de meios para estabelecer um fluxo de controle (interação e testes). Em LEGAL, utilizamos os comandos `FOR...TO...STEP` e `WHILE` para estabelecer os laços e o comando `IF...THEN...ELSE` para testar condições.

Como exemplo, tome-se o caso de cálculo do índice de vegetação normalizado (IVDN) calculado a partir das bandas 1 e 2 do sensor AVHRR do satélite NOAA. Como o índice pretende medir o “stress” hídrico da vegetação, ele precisa ser computado a partir de localizações sem a presença de nuvens. Assim, para o cálculo deste índice, tomamos uma sequência de imagens (pelo menos quinze dias). Se o ponto não estiver coberto por nuvens, calculamos o índice; senão, tenta-se a próxima imagem. O que se espera é que cada ponto esteja livre de nuvens pelo menos uma vez a cada quinze dias.

O programa a seguir supõe ainda a existência de uma representação matricial para os geo-campos das classes `Imagem_NOAA` e `Indice_Vegetacao`.

```

VARIABLE   ivdn               Indice_Vegetacao;
COLLECTION avhrr1, avhrr2  Imagem_NOAA ;
ITERATOR   b1      OVER     avhrr1;
ITERATOR   b2      OVER     avhrr2;
avhrr1= (SELECT imagem FROM imagem IN Imagem_NOAA
        WHERE imagem.banda = 1
        AND imagem.mes = 9
        AND imagem.quinzena = 1);
avhrr2= (SELECT imagem FROM imagem IN Imagem_NOAA
        WHERE imagem.banda = 2
        AND imagem.mes = 9
        AND imagem.quinzena = 1);
ivdn = NEW
        (nome= "IVDN Set95 Quinz1",
        mes = "Setembro", quinzena=1)
        REPRESENTED_BY RASTER
        (RES_X= 1000, RES_Y =1000);

WHILE (MORE (b1))
BEGIN
    IF (VALUE(b1) > 200)
    THEN
    IF (ivdn = 0)
    THEN
        ivdn = (VALUE(b2) - VALUE(b1))
                /(VALUE(b2) + VALUE(b1));
    b1 = NEXT (b1);
    b2 = NEXT (b2);
END

```

O programa acima merece alguns comentários:

1. A recuperação das coleções `avhrr1` e `avhrr2` dá origem a dois conjuntos de geo-campos no formato matricial. Implicitamente, o programa supõe que se dispõe de uma imagem por dia, para todos os dias da quinzena.
2. A variável `ivdn` irá conter o resultado final da operação, na forma de uma matriz de resolução 1000 x 1000 metros (a mesma das imagens AVHRR). Ao ser criada, ela será inicializada com todos os valores em zero.
3. A interação sobre as coleções `avhrr1` e `avhrr2` é feita através dos ponteiros `b1` e `b2`. O operador `VALUE (b1)` devolve o valor do geo-campo apontado pelo ponteiro.
4. A operação é feita de forma implícita sobre as representações matriciais dos geo-campos. Assim, a operação

$$\text{ivdn} = (\text{VALUE}(\text{b2}) - \text{VALUE}(\text{b1})) \\ / (\text{VALUE}(\text{b2}) + \text{VALUE}(\text{b1}));$$

é computada sobre todos os elementos da matriz `ivdn`.

MECANISMOS DE INTEGRAÇÃO TEMPORAL

Os métodos de análise discutidos nas seções anteriores consistem em operações que computam relacionamentos entre objetos e geo-campos, tomados como visões estáticas da realidade geográfica. Deste modo, tais métodos são mais adequados a descrever *padrões estáticos* do que a representar *processos de evolução dinâmica*.

No entanto, a realidade geográfica está em constante evolução, sendo a ação humana um dos principais fatores a engendrar transformações duradouras do espaço. Como afirma Gregory (1992): “é imprescindível que a investigação dos processos seja feita de forma a auxiliar a compreensão da dinâmica atual do ambiente, da maneira de como funcionou no passado e de como poderá operar no futuro.”

Para que o SIG possa funcionar como apoio à modelagem de processos, será preciso dispor de modelos de simulação numérica que descrevam adequadamente os processos a ser estudados (com o possível uso de técnicas de estatística espacial) e integrar estas técnicas com as ferramentas de manuseio, recuperação e apresentação de dados espaciais.

A integração de modelos de simulação e de ferramentas de estatística espacial é vista como uma necessidade fundamental para suplantar as limitações dos métodos de análise da atual geração de SIGs. Para uma discussão em maior detalhe sobre estas questões, veja-se Goodchild et al. (1992), Goodchild et al. (1993), Bregt (1993) e Kemp (1993).

Um problema importante em modelagem ambiental é poder realizar uma integração temporal de geo-campos que expressem fenômenos naturais. Esta necessidade é particularmente premente em problemas de modelagem ambiental. Assim, precisamos incluir na sintaxe de LEGAL provisão para processos de integração.

Para representar o processo de integração, lançamos mão de mecanismos do tipo FOR...TO...STEP.

Para analisar o caso de uma função de integração, considere uma situação onde se quer estimar a aptidão agrícola futura do solo, considerando a manutenção de uma cultura anual. Este é um problema clássico em análise de áreas agrícolas (Burrough, 1986). Possuímos os dados para a situação atual, e desejamos construir *cenários preditivos* para tentar orientar o uso e a ocupação do solo.

Baseado em Burrough (1986), considera-se que a perda de qualidade de solo provocada pela erosão é o principal fator de degradação dos solos e consequente perda de aptidão agrícola. Para estimar esta perda, pode-se utilizar modelos empíricos, como a *equação universal de perda de solos*, que fornece a perda em profundidade do solo como função do tipo de solo, declividade e profundidade atual.

Em nosso exemplo, tomamos um caso simplificado, onde a perda de solo é função apenas da declividade, da profundidade do solo e de um coeficiente de desgaste e a adequação agrícola é computada baseada apenas na profundidade do solo e declividade.

```

// Definicao de variaveis
VARIABLE   prof_solo           Prof_Solo;
VARIABLE   declividade         Declividade;
VARIABLE   perda_solo          Perda_Solo;
VARIABLE   aptidao_hoje        Aptidao;
VARIABLE   aptidao_fut Aptidao;
// Recuperacao de dados
prof_solo:= (SELECT gc FROM Prof_Solo
             WHERE gc.nome = "prof_solo_Quixada");
declividade:= (SELECT gc FROM Declividade
              WHERE gc.nome = "declive_Quixada");
// Instanciacao dos novos geo-campos
perda_solo:= NEW (nome="perda_solo")
              REPRESENTED_BY RASTER (RES_X=100, RES_Y=100);
aptidao_hoje:= NEW (NOME="aptidao_Quixada_1995"
                   REPRESENTED_BY RASTER (RES_X=100, RES_Y=100);
aptidao_fut:= NEW (NOME="aptidao_Quixada_2000",
                  REPRESENTED_BY RASTER (RES_X=100, RES_Y=100);
// Calcule a aptidao hoje
aptidao_hoje:= SWITCH (
    "Boa": prof_solo > 0.30 m E declividade < 10% ;
    "Ruim": OTHERWISE);
// Calcule a profundidade do solo daqui a 5 anos
FOR I = 1 TO 5 STEP 1
BEGIN
    perda_solo:= 0.05*declividade*prof_solo;
    prof_solo:= prof_solo - perda_solo;
END
// Calcule a aptidao estimada futura
aptidao_fut:= SWITCH (
    "Boa": prof_solo > 0.30 m E
           declividade < 10% ;
    "Ruim": OTHERWISE);

```

6.9 INCLUSÃO DE FUNÇÕES EXTERNAS

É rigorosamente impossível prever todo o conjunto de funções a ser utilizado em análise espacial. Deste modo, LEGAL deve embutir previsão para a chamada de funções externas, desenvolvidas pelo usuário e que devem ser ligadas ao sistema.

Por coerência com o paradigma de implementação a ser utilizado (descrito nos próximos capítulos), sugerimos que LEGAL disponha de um *arquivo de descrição*, aonde o desenvolvedor da nova função possa “registrar” sua existência e um *arquivo de configuração*, que relacione o comando em LEGAL com o correspondente programa executável. O arquivo de descrição teria (numa primeira aproximação) uma sintaxe do tipo:

```
COMANDO,
PARAMETRO = VALOR
```

Já o arquivo de configuração deveria incluir o relacionamento entre o comando em LEGAL e o programa externo a ser chamado, como:

```
COMANDO,
programa = <nome do executavel>;
```

Como exemplo, considere-se a existência de uma função de segmentação de imagens, não incluída em LEGAL mas a ser chamada externamente. Sua chamada poderia ser como a mostrada a seguir.

```
EXTERNAL SEGMENTA (BANCO = "Amazonia",
                   PROJETO = "Sao Felix do Xingu",
                   NUMBANDAS = 3, IMAGEM_1="TM3",
                   IMAGEM_2="TM4", IMAGEM_3="TM5",
                   LIMIAR_1=10, LIMIAR_2=8);
```

6.10 LINGUAGEM DE APRESENTAÇÃO

Muitos trabalhos na literatura indicam a importância de permitir o controle de apresentação de forma independente do resultado da operação realizada (veja-se Egenhofer, 1994 e Hemerly, 1993a). A conclusão destes estudos é que se requer uma *linguagem de apresentação* separada da linguagem de consulta.

A linguagem de apresentação apresenta as seguintes necessidades (Egenhofer, 1994; Hemerly, 1993a):

- *Associação de descritores de apresentação gráfica e pictórica aos objetos geográficos*: definição de legendas e modos de apresentação gráfica para os dados geográficos.
- *Controle dos objetos apresentados*: associação de diferentes formas de apresentação (cor, texto, legenda, escala) para diferenciar os resultados da consulta.
- *Combinação dos resultados de consulta*: instrumento bastante poderoso em interfaces de consulta, ao realizar operações de união, intersecção e diferença entre consultas subsequentes.
- *Apresentação do contexto espacial*: a resposta a uma consulta espacial nem sempre pode ser interpretada por si só. Por exemplo, a resposta à consulta "mostre a cidade de São José dos Campos" deve conter ainda a localização da cidade em relação ao mapa do estado de São Paulo.
- *Informação adicional associada*: apresentação de gráficos associados a atributos dos objetos sob forma de cartogramas, barras e discos ("pizza").

Além destes princípios, podemos adicionar o uso de princípios de apresentação que facilitem a percepção cognitiva dos resultados. Neste contexto, o trabalho de Bertin (1973) sobre "variáveis visuais" é de grande importância, bem como os preceitos expressos em Monmonier (1992).

Estes mecanismos serão descritos no que segue.

Descritores gráficos dos objetos geográficos

Para prover mecanismos para controle da visualização dos dados, a linguagem inclui o comando `DEFINE VISUAL`. Os parâmetros controlados pela *definição visual* podem variar de acordo com cada especialização:

- **DTM:** valor máximo e mínimo das isolinhas a ser plotadas, cor e estilo das isolinhas (e das isolinhas-mestras), apresentação (opcional) dos valores da grade, lista (opcional) do valores das isolinhas a ser mostrados.

```
DEFINE VISUAL FOR CLASS <mnt>

    CONTOUR_COLOUR <cor>

    CONTOUR_TEXT <YES | NO >

    PLOT_MODE < NUMBER | INTERVAL >

    CONTOUR_NUMBER <numero>

    CONTOUR_INTERVAL <intervalo>

    GRID_VALUES < YES | NO >

END
```

- **THEMATIC:** aparência das geoclasses (cor, estilo e padrão de preenchimento, espessura das linhas).

```
DEFINE VISUAL

FOR CLASS <tematico> AND THEME <tema>

    COLOUR <cor>

    PATTERN <NONE | SOLID | DOT | HATCH >

    HATCH <angulo de hachura>

    THICKNESS <espessura>

END
```

- `IMAGE`: tabela de cores associada à visualização

```
DEFINE VISUAL
FOR CLASS <imagem>
    COLOUR_TABLE <tab_cor>
END
```

- `GEOOBJECT` : símbolo, texto, cor, estilo de linha e padrão de preenchimento.

```
DEFINE VISUAL
FOR CLASS <objeto>
    COLOUR <cor>
    PATTERN <NONE | SOLID | DASH | HATCH >
    HATCH <angulo de hachura>
    SYMBOL <nome do simbolo>
END
```

Na prática, é mais conveniente construir uma interface interativa para definição dos modos de apresentação padrão associados a cada classe, pois esta escolha não é dinâmica.

Controle dos objetos apresentados

Os princípios enunciados por Bertin (1973) contemplam um sistema de “variáveis visuais” - forma, cor, orientação, textura, valor e tamanho. Bertin argumenta que nossa percepção é realçada ao utilizarmos uma gradação contínua em apenas uma destas variáveis, e que ficamos confusos ao interpretar gráficos aonde estas variáveis se apresentam misturadas.

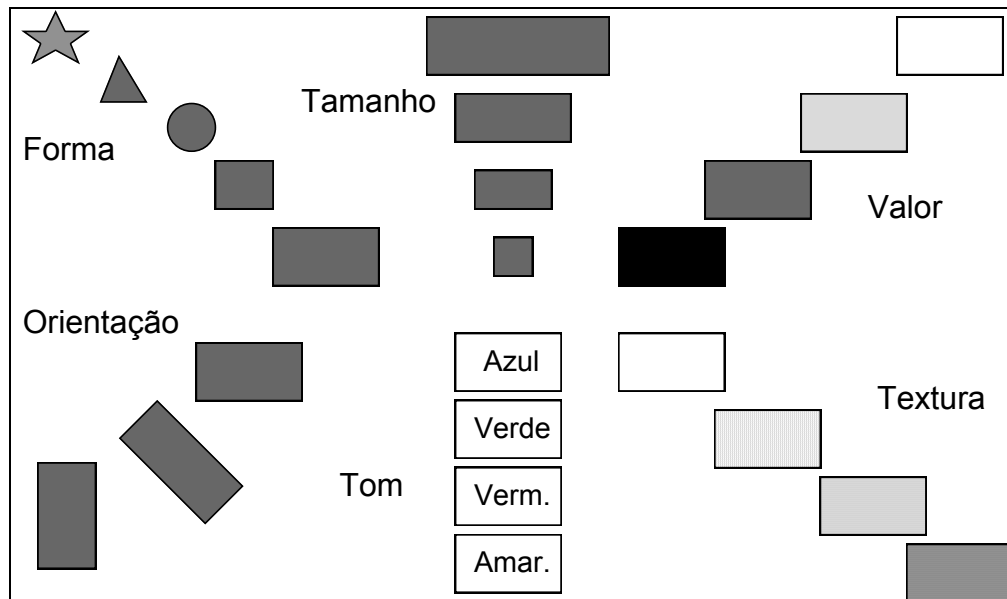


Figura 6.5 - Variáveis visuais de Bertin.

Monmonier (1992), ao discutir o problema de apresentar dados ordinais (valores numéricos referentes a geo-objetos, como população de cidades) alerta para o fato de que o agrupamento destas variáveis pode produzir sérias distorções em sua interpretação. Com base nestes preceitos, utilizamos o comando `SHOW-AS-WHERE-GROUP BY` para controlar a visualização dos geo-objetos selecionados, com a seguinte sintaxe:

```
SHOW <colecão de geo-objetos | valores >
AS < SHADING | SYMBOL | CARTOGRAM | CHART | TEXT >
GROUP BY < VALUES | INTERVALS >
WHERE <parametros>
```

A cláusula `AS` controla o tipo de apresentação, que pode ser:

- `SHADING`: os geo-objetos selecionados são mostrados como polígonos preenchidos.
- `SYMBOL`: associa símbolos aos objetos selecionados.
- `CARTOGRAM`: apresenta os valores de atributos dos objetos sob a forma de círculos cujo área é proporcional ao valor. Pode ser modificado pelos cláusula `GROUP BY`.

- `CHART`: apresenta os valores de atributos de geo-objetos sob forma de gráfico (barras, colunas, areas, pizza, dispersão), podendo ser modificado pela cláusula `GROUP BY`.
- `TEXT`: associa um texto descritivo aos geo-objetos (por omissão, o atributo `NOME` associado ao objeto).

A cláusula `GROUP BY` indica o tipo de agrupamento a ser feito nos valores dos conjuntos de objetos, com as opções de valores ou intervalos. O agrupamento poderá ser feito em escala logarítmica ou linear.

A cláusula `WHERE` indica os parâmetros associados a cada tipo de apresentação, descritos a seguir.

No caso de agrupamento por valores (`GROUP BY VALUES`), podemos ter um agrupamento em dois, três, quatro ou cinco grupos (respectivamente `HALFS`, `THIRDS`, `QUARTILES` e `QUINTILES`), definidos pelo parâmetro:

```
VALUE_MODE= HALFS | THIRDS | QUARTILES | QUINTILES;
```

No caso de agrupamento por intervalos (`GROUP BY INTERVALS`), , podemos ter intervalos igualmente espaçados, com valores máximo e mínimo, e número de intervalos desejado, ou podemos ter intervalos definidos pelo usuário, com os seguintes parâmetros:

```
INTERVAL_MODE = EQUAL | UNEQUAL;
```

```
INTERVAL_MAX  = <valor maximo>;
```

```
INTERVAL_MIN  = <valor minimo>;
```

```
INTERVAL_COUNT = <numero de intervalos>;
```

```
INTERVAL_1    = [<valor minimo>, <valor maximo>;
```

```
INTERVAL_n    = [<valor minimo>, <valor maximo>;
```

O parâmetro `SCALE` permite a escalonamento dos valores do conjunto de objetos a ser agrupado através da opção `GROUP BY`, com a escolha entre as escalas linear e logarítmica. O uso de escala não-linear é importante para evitar distorções na apresentação, como as descritas em Monmonier (1992). A escala de

agrupamento (linear ou logaritmica) será controlada pelo parâmetro `SCALE`, como segue:

```
SCALE = LINEAR | LOGARITHM;
```

No caso das opções `CARTOGRAM` e `CHART`, utilizamos por omissão o agrupamento em quartis e a escala linear.

A apresentação em polígonos sombreados (`SHADING`) pode ser controlada por parâmetros :

- O parâmetro `SHADING_MODE` controla a escolha do mecanismo de sombreamento: polígonos sombreados de uma única cor (`COLOUR`), uma palheta contínua de cores (`PALLETE`), um padrão único (`PATTERN`), ou um conjunto sequencial de hachuras (`HATCH`).

```
SHADING_MODE = < NONE | COLOUR | PALLETE |  
PATTERN | HATCH > (DEFAULT = PATTERN)
```

- A cor do sombreamento, no caso de uma única cor, é controlada pelo parâmetro `SHADING_COLOUR`.

```
SHADING_COLOUR = <cor> (DEFAULT = RED)
```

- Quando da apresentação de uma variável numérica (como população) organizada por grupos, estes podem ser mostrados através de uma palheta contínua, que varia no espaço de tons (considerado como um círculo onde se percorre do azul ao púrpura, passando pelo cian, verde, amarelo, laranja e vermelho). Os pontos inicial e final da palheta são estabelecidos por `SHADING_PALLETE_MIN` e `SHADING_PALLETE_MAX`.

```
SHADING_PALLETE_MIN = <cor> (DEFAULT = BLUE)
```

```
SHADING_PALLETE_MAX = <cor> (DEFAULT = RED)
```

- O uso de um padrão de apresentação é indicado pela variável `SHADING_PATTERN`.

```
SHADING_PATTERN = <padrao>
```

- O uso de hachuras (`SHADING_MODE = HATCH`) é útil para demarcar uma gradação de texturas de apresentação, em conjunto com a cláusula `GROUP BY` e o parâmetro `SCALE`. Supomos uma gradação semelhante à expressa na figura 6.5.

A apresentação de símbolos pode ser controlada pelos parâmetros:

- `SYMBOL_NAME` indica o nome do símbolo.
- `SYMBOL_COLOUR` indica a cor do símbolo.
- `SYMBOL_SIZE` determina o tamanho do símbolo.
- `SYMBOL_ANGLE` denota a orientação do símbolo.

A apresentação dos dados em forma de cartogramas é controlada pelos parâmetro `CARTOGRAM_COLOUR`, que indica a cor dos círculos.

A apresentação dos dados em gráficos (`CHART`) tem as opções:

- O tipo de gráfico, controlado pelo parâmetro `CHART_SHAPE`:

```
CHART_SHAPE = < 2DAREA | 3D AREA | 2D_COLUMNS |
3D_COLUMNS | 2D_BAR | DISPERSION | PIE_CHART >
```

- As cores associadas a cada uma das variáveis apresentadas:

```
CHART_COLOUR_1 = <cor da primeira variavel>
```

```
CHART_COLOUR_2 = <cor da segunda variavel>
```

```
CHART_COLOUR_3 = <cor da terceira variavel>
```

A apresentação de texto (`TEXT`) pode ser controlada a partir dos parâmetros:

- `TEXT_NAME` indica o nome do texto.
- `TEXT_COLOUR` indica a cor do texto.
- `TEXT_SIZE` determina o tamanho do texto.
- `TEXT_ANGLE` denota a orientação do texto.

Apresentação dos resultados de consulta

O comando `SET MODE` controla a combinação dos resultados da consulta com consultas anteriores (Egenhofer, 1990), com a sintaxe:

`SET MODE < NEW | UNION | INTERSECT | REMOVE >`

onde:

- `NEW` indica que a tela deve ser limpa antes de mostrar o resultado da próxima consulta;
- `UNION` combina o resultado da próxima consulta com o resultado da consulta anterior;
- `INTERSECT` seleciona todos os objetos que estão presentes na tela e que pertencem ao resultado da consulta atual;
- `REMOVE` elimina o resultado da consulta atual da apresentação existente.

Exemplo de apresentação em LEGAL

Para ilustrar a linguagem de apresentação, consideremos o caso de um estudo sobre a diversidade regional no Brasil, a partir dos dados de distribuição de miséria (pessoas cujas famílias ganham até 1/2 salário mínimo por mes), com base nos dados do IBGE, de 1985, mostrados na Tabela 6.3.

TABELA 6.3 - PESSOAS COM RENDA FAMILIAR ATÉ 1/2 S.M.

<i>Região</i>	<i>Número (mil)</i>
Norte	1.325
Nordeste	25.830
Sudeste	16.076
Sul	6.551
Centro-Oeste	3.389

Fonte: IBGE (PNAD, 1985).

Inicialmente desejamos apresentar o mapa do Brasil e os dados de miséria sob forma de cartograma. Para tanto iniciamos uma nova consulta e apresentamos o mapa do Brasil (regiões).

```

COLLECTION regioes          Regioes_Brasil;

VARIABLE    mapa_reg_BR    Mapa_Brasil;

SET MODE new;

mapa_reg_BR:= (SELECT mapa FROM mapa IN Mapa_Brasil
                WHERE mapa.nome = "mapa_regioes_BR");

SHOW mapa_reg_BR

    AS SHADING

    WITH SHADING_MODE = NONE;

```

Em cima deste mapa, queremos apresentar um cartograma com a distribuição regional de pobreza. Para isto, escolhemos o modo de união dos resultados da consulta.

```

SET MODE overlay;

regioes:= (SELECT regioes
            FROM regioes IN REGIOES_BRASIL);

SHOW regioes.miseria

    AS CARTOGRAM

    WHERE CARTOGRAM_COLOUR = "GRAY"

    AND SCALE = LINEAR;

```

O resultado da união destas duas consultas é mostrado na figura 6.6.

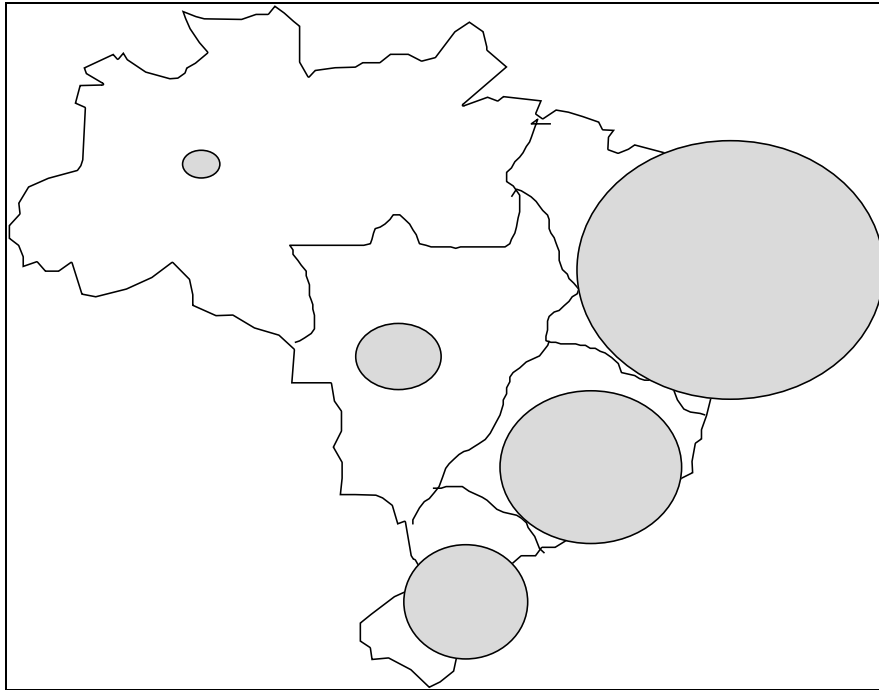


Figura 6.6 - Resultado de combinação de apresentações.

Também podemos querer mostrar estes mesmos dados sob a forma de gráfico:

```
SET MODE NEW;  
SHOW regioes.miseria  
    AS CHART  
    WHERE CHART_MODE = 3D_COLUMN  
    AND SCALE=LINEAR;
```

O gráfico resultante é mostrado na figura 6.7.

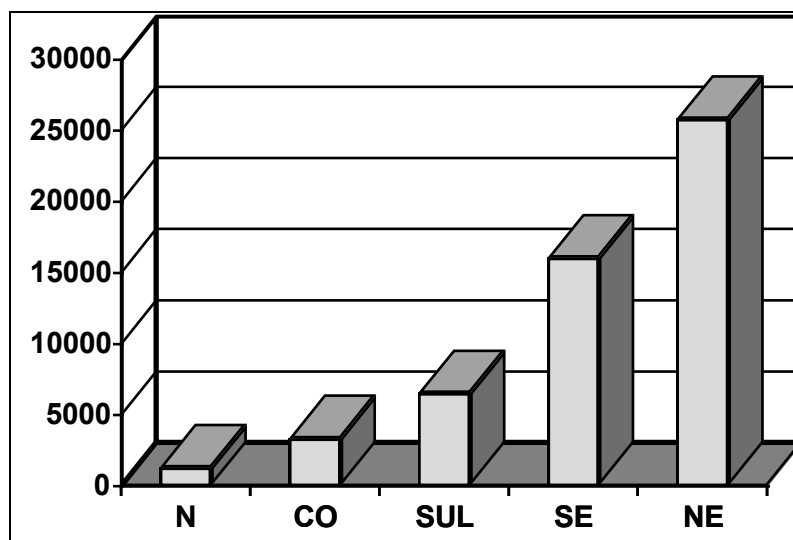


Figura 6.7 - Apresentação em forma de gráfico de barras.

6.1.1 BREVE ANÁLISE COMPARATIVA

Para concluir esta primeira descrição da linguagem LEGAL, é interessante compará-la com duas propostas apresentadas na literatura: SPATIAL SQL (Egenhofer, 1994) e as extensões da linguagem de consulta do SGBD O₂ (OQL) apresentadas em David e Voisard (1993).

Na definição de LEGAL, adotamos o mesmo princípio preconizado em SPATIAL SQL: a separação entre linguagem de consulta e linguagem. Consideramos, no entanto, que SPATIAL SQL ainda carrega as limitações do modelo relacional, o que torna mais complexo o processo de lidar com objetos individuais. O resultado da seleção continua sendo um tupla, e não um conjunto de valores e objetos.

Com relação às extensões de OQL apresentadas em David e Voisard (1993), temos uma situação aonde o modelo de dados utilizado estabelece uma forte ligação entre um geo-objeto e sua representação. Por contraste, o modelo de dados de LEGAL permite uma separação entre os conjuntos de geo-objetos e os mapas cadastrais que contém uma possível representação deste conjunto.

Em resumo, consideramos que LEGAL representa uma avanço com relação às alternativas de linguagem espacial disponíveis na literatura.

6.12 RESUMO DA LINGUAGEM

Apresentamos a seguir um resumo parcial da sintaxe da linguagem LEGAL, em uma notação informal semelhante à BNF. A sintaxe apresentada está baseada na implementação parcial de LEGAL no SPRING-2.0 feita por Ubirajara Freitas e João Pedro Cordeiro e no trabalho de Chu (1994).

- As palavras reservadas de LEGAL (símbolos terminais da gramática) são dados em `COURIER` `MAÍSCULAS`.
- Os símbolos entre `<>` são não-terminais e aqueles precedidos por `@` são reconhecidos pelo analisador léxico.
- O símbolo `<@nome>` indica a sintaxe de identificadores, p.ex: "Uso do Solo"
- O símbolo `<@num>` indica o reconhecimento de um numero inteiro ou na forma decimal.
- O símbolo `[expressão]` indica uma expressão opcional.

Definição de Classes

```
definicao_classe ::=
    CREATE <@nome_classe> (<atributos_conv>)
    IS_A <classe_modelo> | <@nome_classe> [<temas>];
<atributos_conv> ::= [<atributos_conv> ,] <atributo>
<atributo> ::= ATTRIBUTE <@nome_atrib> <classe_atrib>
<temas> ::= THEMES '= {' <atr_tematicos> '}'
<atr_tematicos> ::= [<atr_tematicos> ,] <@nome_tema>
<classe_modelo> ::= THEMATIC | CADASTRAL |
    GEOOBJECT | NUMERIC | IMAGE | NETWORK
<classe_atrib> ::= INT | FLOAT | STRING |
    MULTIMEDIA
```

Organização de Programa

```

<programa LEGAL> ::= { <lista_de_comandos> }
<lista_de_comandos> ::= <comando> |
    <lista_de_comandos> <comando>
<comando> ::= <declaracao> | <instanciacao> |
    <consulta> | <manipulacao> | <transformacao> |
    <apresentacao>

```

Declaração de Variáveis, Coleções e Atributos

```

<declaracao> ::= <declaracao_variavel> |
    <declaracao_colecao> | <declaracao_atributo> |
    <declaracao_tabela>
<declaracao_variavel> ::=
    VARIABLE <@nome_variavel> <@nome_classe>
<declaracao_colecao> ::=
    COLLECTION <@nome_colecao> <@nome_classe>
<declaracao_atributo> ::=
    ATTRIBUTE <@nome_atributos> <classe_atributo>
<declaracao_tabela> ::=
    TABLE <@nome_tabela> SLICE_TABLE |
    WEIGHT_TABLE | RECLASSIFICATION_TABLE

```

Instanciação

```

<instanciacao> ::= <inst_PI> | <inst_tabela>
<inst_PIs> ::=
    <@variavel_def> = NEW(<param_PI>)
    REPRESENTED_BY <geometria> (param_repres)

```

```

<geometria>::= RASTER | VECTOR

<param_PI>::= <param_PI> | <parametro>

<parametro>::= <@nome_atrib> = <@valor_atrib>

<param_repres>::=
    RESX = <@num>,
    RESY = <@num>,
    SCALE = <@num>,
    BOX= [<@lat1>, <@long1>, <@lat2>, <@long1>]

<inst_tabela>::= <@nome_tabela>
    = NEW (<expr_cria_tabela>)

<expr_cria_tabela>:=
    CLASS_IN = <@nome_classe>,
    TYPE = WEIGHT, <lista_de_pesos> |
    CLASS_OUT = <@nome_classe>,
    TYPE = SLICE, <lista_de_fatias> |
    CLASS_IN = <@nome_classe>,
    CLASS_OUT = <@nome_classe>,
    TYPE = RECLASSIFY, <lista_de_classes>

<lista_de_pesos>::= <atrib_peso> |
    <lista_de_pesos> , <atrib_de_peso>

<atrib_peso> ::= <@tema> : <@num>

<lista_de_fatias>::= <atrib_fatia> |
    <lista_de_fatias> , <atrib_fatia>

<atrib_fatia>::= <@tema> : <@num> , <@num>

<lista_de_classes>::= <atrib_classe> |
    <lista_de_classe> , <atrib_classe>

```

<atrib_classe>::= <@tema_saida> : <@tema_entrada>

Consulta Espacial

consulta_LEGAL::= <@nome_colecao> = <expr_consulta>

<expr_consulta>::=

SELECT <objetos_e_valores>

FROM <@nome_objeto> IN <@nome_classe>

[ON MAP <@nome_cadastral> | <@nome_rede>]

[<WHERE_clause>]

<objetos_e_valores>::= <objetos> | <valores>

<objetos>::= <objetos> | <@nome_objeto>

<valores>::= <valores> | <@nome_atributo>

<WHERE_clause>::= WHERE <boolean>

<boolean>:= <bool_expr1>

| <boolean> OR <bool_expr1>

<bool_expr1>::= <bool_expr1> AND <bool_expr2> |

<bool_expr2>

<bool_expr2>::= [NOT] <bool_prim>

<bool_prim>::= <predic_topologico> | <predic_metrico>

| <pred_aritm> | <boolean>

<predic_topologico>::=

<@nome_geoobjeto> <rel_topologia>

<@nome_geoobjeto> | <regiao_tematica>

<rel_topologia>::= INSIDE | OVERLAP | DISJOINT |

CROSS | TOUCH | LINKED_TO

<regiao_tematica>::=

REGION(<@nome_geocampo_tematico> == <@tema>)

<predic_metrico>::=

```

DISTANCE    (<@nome_geoobjeto>,    <@nome_geoobjeto>)
            <op_comp> <@numero>

<predic_aritm> ::= <expr_num> <oper_comp> <expr_num>

<expr_num> ::= <termo_aritm> |
<expr_num> <oper_add> <termo_aritm>

<termo_aritm> ::= <aritml> |
<termo_aritm> <op_mul> <aritml>

<aritml> ::= [op_add] <primario>

<primario> ::= <@num> | <objetos_e_valores>

<op_comp> ::= > | < | >= | <= | = | !=

<op_add> ::= + | -

<op_mul> ::= * | /

```

Manipulação Espacial

```

<manipulacao>::=
    <@nome_geocampo_tematico> = <oper_tematica> |
    <@nome_geocampo_numerico> = <oper_numerica>

<oper_tematica>::=
    <@nome_geocampo_tematico> |
    SLICE ( <oper_numerica>, <@nome_tabela_fatia> ) |
    RECLASSIFY( <oper_tematica>, <@nome_tabela_reclass> ) |
    SWITCH ( { <lista_de_regras> } )

<lista_regras>::= <regra> | <lista_regras>; <regra>

<regra>::= <@tema> : <expr_bool>

<oper_numerica>::= <@numero> |
    <@nome_geocampo_numerico> |

```

```

WEIGHT (<oper_tematica>, <@nome_tabela_pond>) |
<op_local_num> |
<op_zonal> |
FUZZYL(<@nome_geocampo_num>, <@num>, <@num>) |
FUZZYU(<@nome_geocampo_num>, <@num>, <@num>) |
REFINE (<@nome_geocampo_num>, <proc_refin>) |
SLOPE (<@nome_geocampo_num>) |
ASPECT (<@nome_geocampo_num>) |
    <funcao_matematica> (<oper_numerica>) |
    <oper_numerica> |
    <oper_numerica> <op_add> <oper_numerica> |
    <oper_numerica> <op_mul> <oper_numerica> |
    - <oper_numerica>
<funcao_matematica> := SIN | COS | ATAN | LOG |
    LOG10 | EXP | SQRT | INT | ABS
<expr_bool> ::= <expr_bool> && <expr_bool> |
    <expr_bool> || <expr_bool> |
    [<expr_bool>] |
    <@nome_geocampo_tematico>.THEME == <@tema> |
    <@nome_geocampo_tematico>.THEME != <@tema> |
    <expr_numerica> <op_comp> <expr_numerica>
<op_local_num> ::=
    LOCAL_SUM (<oper_numerica>, <viz>) |
    LOCAL_MEAN (<oper_numerica>, <viz>) |
    LOCAL_MAX(<oper_numerica>, <viz>) |
    LOCAL_MIN (<oper_numerica>, <viz>) |

```



```

        LOCAL_STDEV (<oper_numerica>, <viz>)
<viz>::= <retangulo> | <circulo>
<retangulo>::= RECTANGLE, <@num>, <@num>
<circulo>::= CIRCLE, <@num>
<op_zonal>::=
        ZONAL_SUM (<oper_numerica>, <@tematico>) |
        ZONAL_MEAN (<oper_numerica>, <@tematico>) |
        ZONAL_MAX(<oper_numerica>, <@tematico>) |
        ZONAL_MIN (<oper_numerica>, <@tematico>) |
        ZONAL_STDEV (<oper_numerica>, <@tematico>)

```

Transformação entre Campos e Objetos

```

<transformacao>::= <campos_em_objetos> |
        <objetos_em_campos>
<campos_em_objetos>::= <identificacao> |
        <interseccao>
<objetos_em_campos>::= <mapa_distancia> |
        <recl_atributos>
<identificacao>::= <id_cadastral> | <id_objetos>
<id_cadastral>::= <@nome_cadastral> =
        IDENTIFY_MAP (<@nome_tematico>)
<id_objetos>::= <@nome_colecao> = IDENTIFY_OBJECTS
(<par_tema_atr>) ON MAP <@nome_cadastral>
<interseccao>::= <int_cadastral> | <int_objetos>
<int_cadastral>::= <@nome_cadastral> =
        INTERSECT_MAP (<lista_temat>)
<lista_temat>::= <@nome_tematico> | <lista_temat>

```

```

<int_objetos>::= <@nome_colecao> = IDENTIFY_OBJECTS
(<lista_pares>) ON MAP <@nome_cadastral>

<lista_pares>::= <par_tema_atr> | <lista_pares>

<par_tema_atr>::= (<@nome_tematico>,<@nome_atributo>)

<mapa_distancia>::= <@nome_numerico> =
DISTANCE_MAP (<geo_objetos>)

<geo_objetos>::= <@nome_geoobjeto> | <geo_objetos>

<recl_atributo>::= <@nome_geocampo> =
RECL_ATTR (<geo_objetos>, <@nome_atributo>)      ON
MAP <@nome_cadastral>

```

Apresentação

```

<apresentacao>::= <op_mostrar> | <controle_tela>

<controle_tela>::= SET MODE <modo_vis>

<modo_vis>::= NEW | UNION | INTERSECT | REMOVE

<op_mostrar>::= SHOW <objetos_e_valores>

      AS <modo_apresent>

      GROUP_BY <modo_agrupar>

      WHERE <lista_param_apresen>

<modo_apresent>::= SHADING | SYMBOL | CARTOGRAM |
      CHART | TEXT

<modo_agrupar>::= VALUES | INTERVAL

<lista_param_apresen>::= <param_apres> |
      <lista_param_apresen>

<param_apres>::=

      <grup_valores> |

      <grup_intervalos> |

```

```

    <escala>      |
    <sombreamento> |
    <simbolos>   |
    <cartograma> |
    <grafico>    |
    <texto>

<grup_valores> ::= VALUE_MODE = <param_valor>

<param_valor> ::= HALFS | THIRDS | QUARTILES |
    QUINTILES;

<grup_intervalos> ::=
    INTERVAL_MODE = <modo_int> |
    INTERVAL_MAX  = <@num> |
    INTERVAL_MIN  = <@num> |
    INTERVAL_COUNT = <@num> |
    <lista_intervalos>

<lista_intervalos> ::= <intervalo> |
    <lista_intervalos>

<intervalo> ::= INTERVAL.<@num> = <@num>, <@num>

<escala> ::= SCALE = <param_escala>

<param_escala> ::= LINEAR | LOG

<sombreamento> ::=
    SHADING_MODE = <modo_sombr> |
    SHADING_COLOUR = <@nome_cor> |
    SHADING_PALLETE_MIN = <@nome_cor> |
    SHADING_PALLETE_MAX = <@nome_cor> |

```

```

    SHADING_PATTERN = <@nome_padrao>

<modo_sombr>::= NONE | COLOUR | PALLETE |
    PATTERN | HATCH

<simbolos>::=

    SYMBOL_NAME = <@nome_simbolo> |
    SYMBOL_COLOUR = <@nome_cor> |
    SYMBOL_SIZE = <@num> |
    SYMBOL_ANGLE = <@num>

<cartograma>::=

    CARTOGRAM_COLOUR = <@nome_cor>

<grafico>::=

    CHART_SHAPE = <forma_graf>

<forma_graf>::= 2D_AREA | 3D_AREA | 2D_COLUMNS |
    3D_COLUMNS | 2D_BAR | DISPERSION | PIE_CHART

```

7 ARQUITETURAS PARA BANCOS DE DADOS GEOGRÁFICOS

“Quid est corporis pulchritudo? Congruentia partium cum quadam coloris suavitate.”

(O que é a beleza do corpo ? É a proporção das partes acompanhada por uma certa doçura de colorido).

Sto. Agostinho

7.1 VISÃO GERAL

Nos três capítulos anteriores, analisamos os bancos de dados geográficos numa visão externa, ao discutir as questões de modelo de dados, operações e linguagens. Neste capítulo e no próximo, estaremos preocupados com questões de arquitetura de sistemas e de implementação.

Este capítulo analisa as arquiteturas de bancos de dados geográficos. A seção 7.2 apresenta uma breve revisão da literatura sobre o assunto. Na seção 7.3, descrevemos um modelo de referência para bancos de dados geográficos, que separa as funções de armazenamento, manipulação e visualização. Na seção 7.4, discutimos as diferentes estratégias para implementação da arquitetura de referência em ambientes cliente-servidor e classificamos, de maneira coerente, os SIGs disponíveis comercialmente ou propostos como protótipos. A seção 7.5 apresenta uma discussão sobre a implementação da arquitetura de referência num ambiente distribuído.

7.2 BREVE REVISÃO DA LITERATURA

Uma revisão geral sobre o problema de arquiteturas em SIGs é apresentada em Pissinou et al. (1993) e Casanova et. al. (1993). A literatura recente apresenta um significativo conjunto de projetos de pesquisa, que utilizam tecnologias emergentes de bancos de dados (Cattell, 1991) para projetar sistemas de gerência de bancos de dados geográficos. Destes projetos, cabe mencionar:

- GEO++ (Vijbrief and Oosterom, 1992), desenvolvido sobre o SGBD pós-relacional POSTGRES (Stonebraker and Kemnitz, 1991). O sistema POSTGRES é uma extensão ampla da tecnologia relacional e dispõe de

algumas ferramentas para manuseio de dados gráficos, incluindo o suporte para *árvores-R*, estrutura de indexação espacial com bom desempenho e tipos de dados *ponto*, *segmento de linha*, *polilinhas* e *retângulo envolvente* ("*box*"). O sistema GEO++ é escrito em C++ e estende os tipos de dados geométricos do POSTGRES, além de rotinas de acesso aos dados.

- PARADISE (DeWitt et al., 1993; DeWitt et al., 1994) - O projeto PARADISE (**Parallel Data Information System**) tem o objetivo de aplicar as tecnologias de orientação-a-objetos e bancos de dados paralelos para desenvolver um SIG capaz de gerenciar grandes bases de dados espaciais. Como gerenciador persistente de objetos, PARADISE utiliza SHORE (Carey et al., 1994), também desenvolvido pelo grupo da Universidade de Wisconsin.
- SEQUOIA 2000 - O projeto SEQUOIA 2000 (Stonebraker and Dozier, 1992), é um projeto de longo prazo com o propósito de dar suporte a estudos de mudanças globais. O projeto está baseado no SGBD POSTGRES e envolve ainda aspectos de armazenamento, redes e visualização (Gardels, 1993).
- GODOT - (**Geographic Data Management with Object-oriented Techniques**): protótipo desenvolvido pela Universidade de Ulm (Alemanha) que utiliza o sistema OBJECTSTORE, adicionando classes e métodos (Günther and Riekert, 1993);
- GEO₂: desenvolvido pelo Instituto Geográfico Nacional da França com suporte do O₂. (David et al., 1993). O modelo de dados utilizado, distingue entre a semântica dos dados geográficos (atributos e localização) e sua representação (polígonos, redes e mapas). Os autores apontam alguns problemas do SGBD O₂ neste contexto¹⁶.

¹⁶De acordo com informações prestadas pela prof. Cláudia Bauzer Medeiros, o SGBD O₂ está sendo estendido pela O₂ Technology para incluir métodos de acesso para dados espaciais.

- Milne *et al.* (1993) discutem a construção de um SIG baseado no sistema ONTOS. Os autores utilizam o modelo de dados SDTS (Spatial Data Transfer Standard) como base para definição das classes de objetos espaciais. O SGBD ONTOS provê uma biblioteca de classes em C++, que os autores estendem com uma interface gráfica, apresentação de mapas e construção de um banco de dados geográfico. Os autores comparam o desempenho do sistema ao recuperar objetos complexos, com os ambientes SIRO_DBMS (Abel, 1989) e ORACLE, e indicam um ganho significativo de desempenho.

No caso de centros de dados ambientais, o principal conjunto de trabalhos são os estudos que estão sendo realizados para o projeto EOSDIS (Earth Observation System Data and Information System). O EOSDIS representa o mais ambicioso sistema de informação ambiental já construído pela NASA e seu objetivo é prover a comunidade internacional de cientistas com o mais completo acervo de dados disponível sobre a Terra (Asrar and Dozier, 1994).

7.3 UMA ARQUITETURA DE REFERÊNCIA PARA GERÊNCIA DE OBJETOS GEOGRÁFICOS

Para tratar de todos os tipos de dados geográficos e em função de requisitos gerais como os apresentados no capítulo 3, Hemerly et al. (1994) propõem um arquitetura em camadas para sistemas de gerência de bancos de dados geográficos, com uma visão abrangente do problema. A arquitetura introduz subsistemas especializados para cada tipo de dado geográfico e está apresentada na figura 7.1.

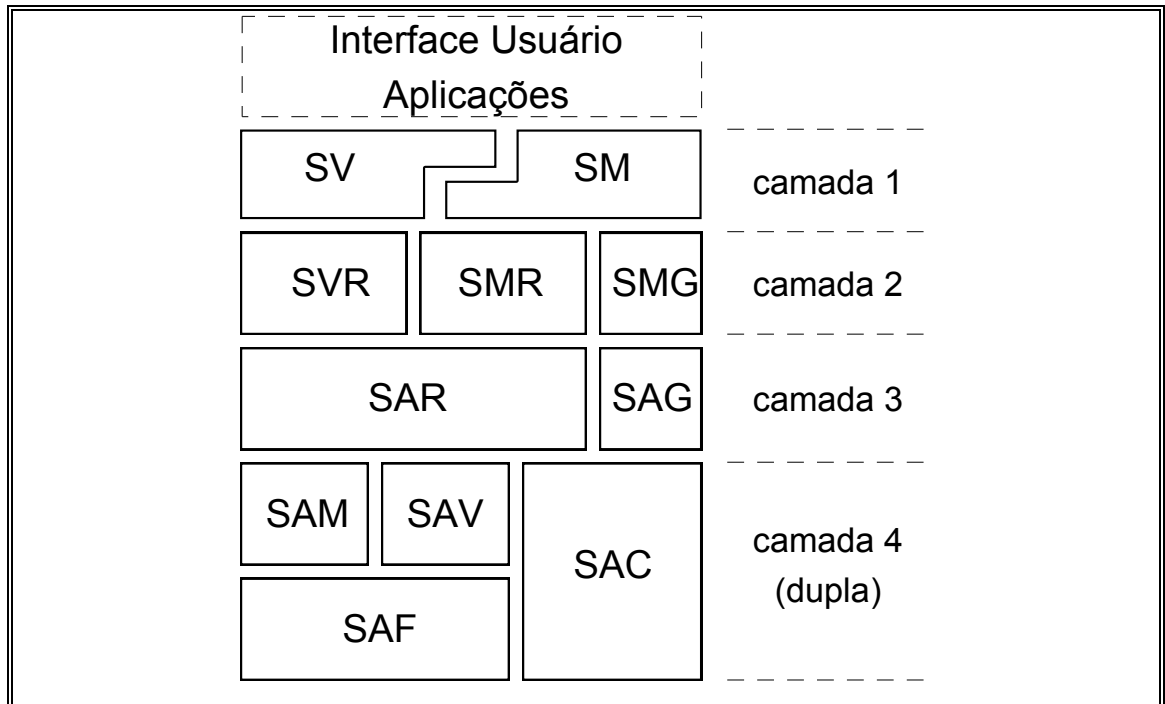


Figura 7.1 - Arquitetura de referência para Gerência de Dados Geográficos

A descrição destas camadas é feita em Tocantins e Casanova (1995), como segue. No nível superior, não especificado pela arquitetura, encontram-se as aplicações de geoprocessamento e módulos de interface com o usuário.

A primeira camada separa as questões de visualização das funções de manipulação e oferece as abstrações correspondentes a *geo-objetos*, *geo-campos* e *mapas de geo-objetos* (conforme o modelo de dados apresentado no capítulo 4), além de métodos para manipular e visualizar estes tipos de dados e suas extensões. O *subsistema de visualização* (SV) oferece funções básicas para apresentação de objetos, tipicamente mais complexas que as necessárias para dados convencionais. O *sub-sistema de manipulação* (SM) oferece funções para definição e manipulação de *geo-objetos*, *geo-campos*, *mapas de geo-objetos* e suas representações, procurando esconder do usuário os problemas de representação geométrica (matrizes x vetores).

A segunda camada desce um degrau na escala de abstração, e separa os componentes espacial e convencional dos dados geográficos. Oferece serviços de visualização e manipulação de alto nível para representações geométricas, através do *sub-sistema de visualização de representações* (SVR) e do *sub-sistema de*

manipulação de representações (SMR), e serviços de manipulação dos atributos convencionais dos geo-objetos e de geo-campos, através do *sub-sistema de manipulação de geo-objetos e geo-campos* (SMG).

A terceira camada cobre serviços de armazenamento e manipulação elementar para representações, através do *sub-sistema de armazenamento de representações* (SAR), e para atributos convencionais, através do *sub-sistema de armazenamento de geo-objetos* (SAG). Estes sistemas são responsáveis pelo armazenamento persistente e edição elementar desses tipos de objetos.

A quarta camada implementa as formas de armazenamento - matricial, vetorial e convencional através do *sub-sistema de armazenamento matricial* (SAM), *sub-sistema de armazenamento vetorial* (SAV) e *sub-sistema de armazenamento convencional* (SAC). Os dois primeiros subsistemas fazem uso do serviço de armazenamento de páginas físicas disponível no *sub-sistema de armazenamento físico* (SAF).

O restante deste capítulo discute alternativas para implementação desta arquitetura, com análise dos ambientes cliente-servidor (seção 7.4) e distribuídos (seção 7.5). No que segue, o leitor deverá ter sempre em mente a figura 7.1 e as abreviaturas utilizadas.

7.4 IMPLEMENTAÇÃO EM AMBIENTES CLIENTE-SERVIDOR

Apresentamos nesta seção diferentes estratégias de implementação para a arquitetura de referência, baseadas em sistemas de gerência de bancos de dados com graus crescentes de funcionalidade. A taxonomia proposta nos permite classificar, de maneira coerente, vários SIG disponíveis comercialmente ou propostos como protótipos.

As estratégias de implementação aqui discutidas utilizam um servidor único, responsável pelo armazenamento e recuperação dos atributos convencionais, e diferem essencialmente no tratamento da parte espacial dos objetos geográficos. A classificação apresentada está baseada no trabalho de Vijbrief e vanOosterom (1992).

ESTRATÉGIA DUAL

Em linhas gerais, um SIG implementado com a estratégia dual utiliza um SGBD relacional para armazenar os atributos convencionais dos objetos geográficos (na forma de tabelas) e arquivos para guardar as representações geométricas destes objetos. Um *identificador comum* liga os componentes geométrico e convencional do objeto geográfico. Para recuperar um objeto, os dois subsistemas devem ser pesquisados e a resposta é uma composição de resultados. Esta estratégia é ilustrada na figura 7.2.

Em termos da arquitetura de referência, idealmente esta estratégia estaria limitada a implementar o SAC sobre um SGBD relacional e o SAM e o SAV sobre o sistema de arquivos do sistema operacional. A maior parte das soluções disponíveis comercialmente para Geoprocessamento utiliza esta abordagem e os produtos disponíveis levam a dualidade até as camadas superiores. Exemplos de sistemas comerciais baseados em estratégia dual são o ARC/INFO (Morehouse, 1992), MGE (Intergraph, 1990) e o SPRING (Câmara et al., 1993)¹⁷.

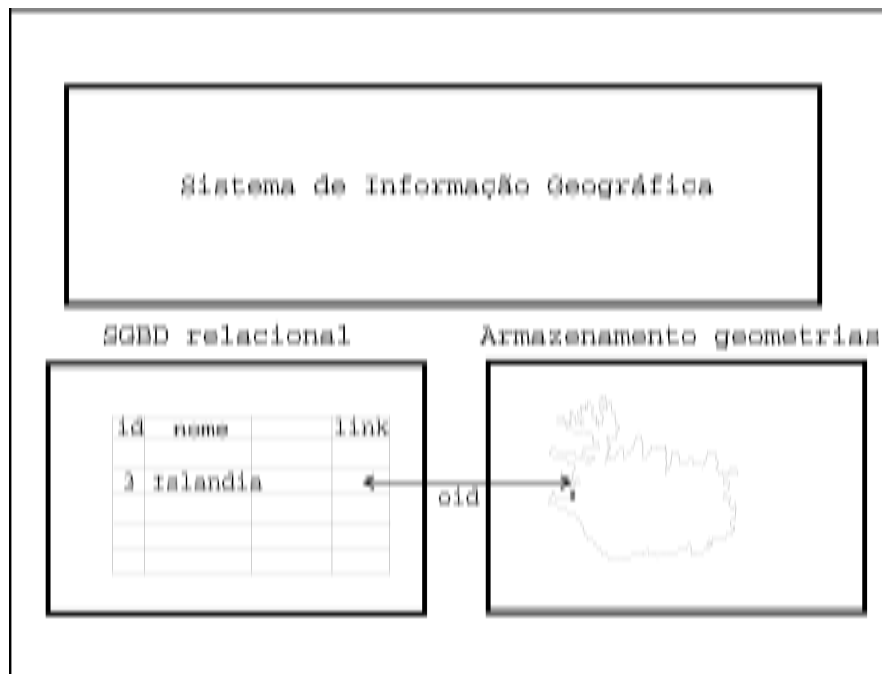


Figura 7.2 - Estratégia dual (adaptado de Vijbrief e vanOosterom (1992)).

¹⁷O SPRING procura esconder do usuário tal dualidade, por meio de seu modelo de dados.

A principal vantagem desta estratégia é poder utilizar os SGBDs relacionais de mercado. No entanto, como as representações geométricas dos objetos espaciais estão fora do controle do SGBD, esta estrutura dificulta o equacionamento das questões de otimização de consultas, gerência de transações e controle de integridade e de concorrência. Estes problemas só podem ser resolvidos através de implementações sofisticadas das camadas superiores da arquitetura genérica, que operem coordenadamente com o SGBD convencional.

Esta alternativa exige ainda a construção do SAM e do SAV apenas a partir dos construtores oferecidos pelo sistema operacional, em geral muito mais pobres que os oferecidos pelo SGBD.

A execução de uma consulta em um ambiente dual tipicamente segue um plano rígido: a parte convencional da consulta é processada pelo SGBD relacional, em separado das restrições espaciais, que são deixadas a cargo de subsistemas fora do SGBD. Tal rigidez na geração do plano de execução, embora conveniente para a implementação do SIG, geralmente leva a planos sub-ótimos.

Adicionalmente, é possível que a indicação da representação geométrica do objeto geográfico (tipicamente um nome de arquivo) permaneça registrada no banco de dados convencional, embora o arquivo contendo o componente geométrico deste objeto tenha sido removido (inadvertidamente ou por interrupção indevida do processamento).

Para resolver completamente o problema de controle da integridade, é necessário implementar as camadas inferiores da arquitetura genérica de forma inacessível ao sistema operacional. Por exemplo, em UNIX, os dados teriam de ser colocados numa “raw partition” (partição da qual o sistema operacional não conhece a estrutura interna), deixando para o SAF a tarefa de guardar informações como localizações de blocos e trilhas de disco, aonde estão armazenadas as representações geométricas.

ESTRATÉGIA BASEADA EM CAMPOS LONGOS

Esta estratégia baseia-se no uso de SGBDs relacionais com suporte para campos longos (um *campo longo* é uma cadeia binária aonde se pode armazenar

informação gráfica, numérica ou pictórica). Fisicamente, o SGBD permite que um campo longo possa ter tamanho significativo (em alguns casos, da ordem de gigabytes), para que possa armazenar as representações geométricas associadas aos objetos geográficos, conforme ilustrado na figura 7.3.

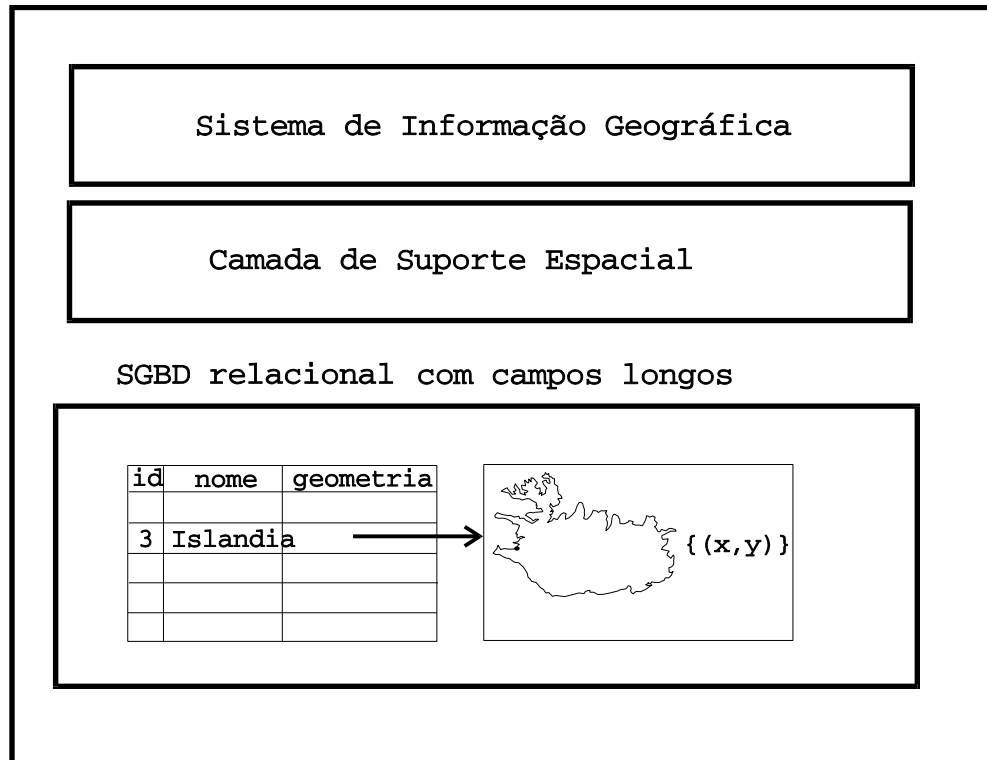


Figura 7.2 - Estratégia baseada em campos longos (adaptado de Vijbrief e vanOosterom (1992)).

As funções oferecidas por um SGBD para tratamento de campos longos são tipicamente limitadas ao acesso a partes do conteúdo do campo, especificando a posição de início e o tamanho do fragmento. Em alguns casos, possuem mecanismos de proteção de integridade física (usualmente *shadowing*).

Em termos da arquitetura de referência, tanto o SAC como o SAM e o SAV passam a ser implementados pelo SGBD relacional, pelo menos em parte. O acesso interno ao conteúdo dos campos longos é deixado a cargo de uma camada implementada sobre o SGBD.

A principal vantagem desta estratégia está em armazenar todos os objetos de um banco de dados geográfico através de um único SGBD, evitando os problemas

de gerência de transações e de controle de integridade e de concorrência, típicos do ambiente dual. No entanto, como o SGBD trata um campo longo como uma cadeia binária, não conhece a semântica do seu conteúdo e, portanto, não possui mecanismos satisfatórios para seu tratamento. Deste modo, a implementação de um SIG utilizando esta estratégia requer que sejam incluídos métodos de acesso espacial e um otimizador de consultas que prepare planos de execução de acordo com tais métodos.

Adicionalmente, embora o padrão SQL, principal interface oferecida pelos SGBD relacionais, defina funções básicas para tratamento de campos longos, há variações nos mecanismos de fato disponíveis nas implementações comerciais. Como consequência, a implementação desta arquitetura pode ficar dependente do SGBD escolhido.

A estratégia de uso dos campos longos é mais freqüente como suporte a aplicações de redes ("facilities management"), aonde a maior parte das consultas é resolvida baseando-se apenas nos atributos não-espaciais. Neste caso, a parte gráfica dos dados é limitada, o que abrandando os requisitos de métodos eficientes de acesso aos dados espaciais. Um dos primeiros exemplos desta estratégia foi o protótipo SIRO-DBMS (Abel, 1989), construído por sobre o SGBD ORACLE. Os produtos comerciais que baseados em campos longos incluem o sistema SYSTEM 9 (van Eck e Uffer, 1989).

ESTRATÉGIA BASEADA EM MECANISMOS DE EXTENSÃO

Esta estratégia depende da adoção de um SGBD que disponha de mecanismos que permitam implementar os componentes de manuseio de dados espaciais através de extensões ao seu próprio ambiente. Exemplos de SGBD extensíveis são os sistemas chamados de pós-relacionais, como o POSTGRES (Stonebraker e Rowe, 1991) e o STARBURST (Haas e Cody, 1991) e orientados a objetos, como o O₂ (Bancilhon et al., 1991) e o OBJECTSTORE.

Através de sua interface de definição de dados, um SGBD extensível tipicamente permite a definição de novos tipos de objetos, como os utilizados em SIGs (esta estratégia é ilustrada na figura 7.4). A especificação de um novo tipo de dados inclui a definição dos atributos dos objetos e dos métodos associados. Este mecanismo é muito mais amplo que o uso de campos longos, pois permite capturar a semântica dos objetos.

Uma das dificuldades inicialmente encontradas para desenvolver aplicações sobre SGBD extensíveis foi a falta de uma padronização de mecanismos para definição e consulta. Reconhecendo esta necessidade, diversos fabricantes de SGBD orientados-a-objetos (incluindo companhias como Object Design, O₂ Technology, Objectivity e Ontos) formaram o ODMG (Object Data Management Group), que publicou o padrão ODMG-93 (Cattell, 1994). Este padrão inclui uma definição de uma linguagem de definição de objetos (ODL - *object definition language*) e de uma linguagem de consulta a objetos (OQL - *object query language*); as empresas participantes do ODMG comprometeram-se a adotar tais padrões a partir de 1995.

Dado o poder expressivo da nova geração de SGBD, há uma grande número de projetos de pesquisa que utilizam esta estratégia para implementar sistemas de Geoprocessamento, como apresentado na seção 7.2. Das implementações industriais, merecem menção os produtos SMALLWORLD (Batty, 1993), que utiliza seu próprio SGBD, e GEO₂ (David et al., 1993), que utiliza o SGBD O₂.

A grande questão em aberto com relação ao uso SGBDs extensíveis para Geoprocessamento reside em seu desempenho ao lidar com grandes volumes de

dados. Para permitir a implementação eficiente de um SIG, o SGBD deveria estar aberto a extensões também em suas camadas mais internas e deveria permitir a definição de novos métodos de armazenamento e acesso, a adição de novas regras ao otimizador de consultas (para fazer bom uso dos novos métodos) e até a modificação da política de gerência de transações para adaptar o SGBD às necessidades das aplicações em Geoprocessamento. Em última instância, deveria ser possível implementar todos os subsistemas da arquitetura de referência como extensões do SGBD. Esta estratégia apresenta dificuldades técnicas, mas aumenta muito a funcionalidade e possivelmente o desempenho da aplicação resultante (Milne et al., 1993).

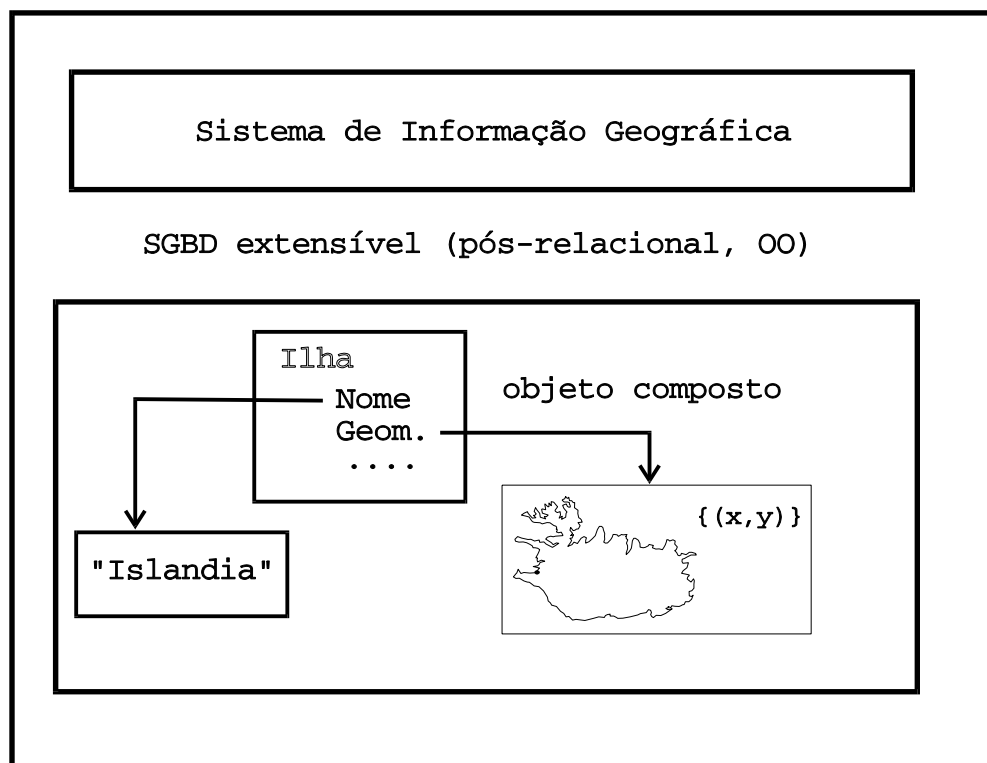


Figura 7.4 - Estratégia baseada em mecanismos de extensão.

RESUMO COMPARATIVO

Uma comparação das vantagens e desvantagens de cada tipo de arquitetura é apresentada na Tabela 7.1. Vale notar que a taxonomia apresentada é esquemática e esconde nuances: as vantagens completas da *arquitetura extensível* só ocorrem quando o SGBD utilizado for completamente extensível.

TABELA 7.1

*COMPARAÇÃO ENTRE ESTRATÉGIAS PARA
BANCOS DE DADOS GEOGRÁFICOS EM AMBIENTE CLIENTE-SERVIDOR*

<i>Estratégia</i>	<i>Vantagens</i>	<i>Problemas</i>
DUAL	1. Uso de SGBD relacionais de mercado 2. Independência de fabricantes de SGBD	1. Controle de integridade 2. Otimização de consultas
CAMPOS LONGOS	1. Controle de integridade 2. Uso de SGBD relacionais de mercado (com campos longos)	1. Dependência de fabricantes de SGBD 2. Otimização de consultas
EXTENSÍVEL	1. Controle de integridade 2. Otimização de consultas e métodos de acesso	1. Desempenho em grandes volumes de dados 2. Falta de compatibilidade com bancos de dados existentes

7.5 ARQUITETURAS DISTRIBUÍDAS

VISÃO GERAL

As arquiteturas distribuídas são uma importante alternativa para a organização tradicional de bancos de dados baseada na abordagem *cliente-servidor*.

Numa arquitetura cliente-servidor tradicional, a máquina cliente recebe comandos do usuário e executa aplicações, que requisitam dados pela rede de uma máquina servidora, tipicamente através de diretivas SQL. O processamento de consultas é realizado no servidor e os resultados enviados para a máquina cliente.

Com o grande aumento da capacidade computacional das estações de trabalho, a distinção entre os desempenhos dos ambientes cliente e servidor

diminuiu bastante. Deste modo, é importante distribuir o processamento ao longo da rede, para aproveitar todo o potencial de uma rede de computadores.

Por esta razão, abordaremos nesta seção diferentes formas de implementação distribuída para SIGs, partindo da arquitetura em camadas da seção 7.3.

ARQUITETURAS ORIENTADAS A SERVIÇOS - INTRODUÇÃO

Numa arquitetura orientada a serviços, a hierarquia cliente-servidor tradicional é quebrada em três componentes: *clientes*, *provedores de serviços* e a *malha de interoperabilidade* (cf. Figura 7.5). A idéia é que os provedores de serviços são especializados em tarefas específicas, e que os diferentes componentes da arquitetura podem estar distribuídos em diferentes máquinas e até diferentes locais. Cada servidor tem apenas a informação necessária para sua operação, sem sobrecarga desnecessária. Os sistemas EOSDIS (Elkington et al., 1994) e METVIEW (Raoult et al., 1995) utilizam arquiteturas orientadas-a-serviços.

O coração desta arquitetura é um *malha de interoperabilidade* (“*request broker*”), responsável pela comunicação entre os clientes e os servidores e que implementa um *modelo de objetos*, uma representação dos objetos que transcende os limites de um único aplicativo e uma única linguagem. Esta malha realiza vários tipos de operações:

- anúncio dos serviços disponíveis para os clientes.
- adequação entre o pedido dos usuários e os serviços disponíveis. Em muitos casos, uma solicitação do usuário pode envolver vários serviços disponíveis na rede.
- tradução e transferência de parâmetros e objetos entre máquinas, processos e espaços de endereçamento distintos.

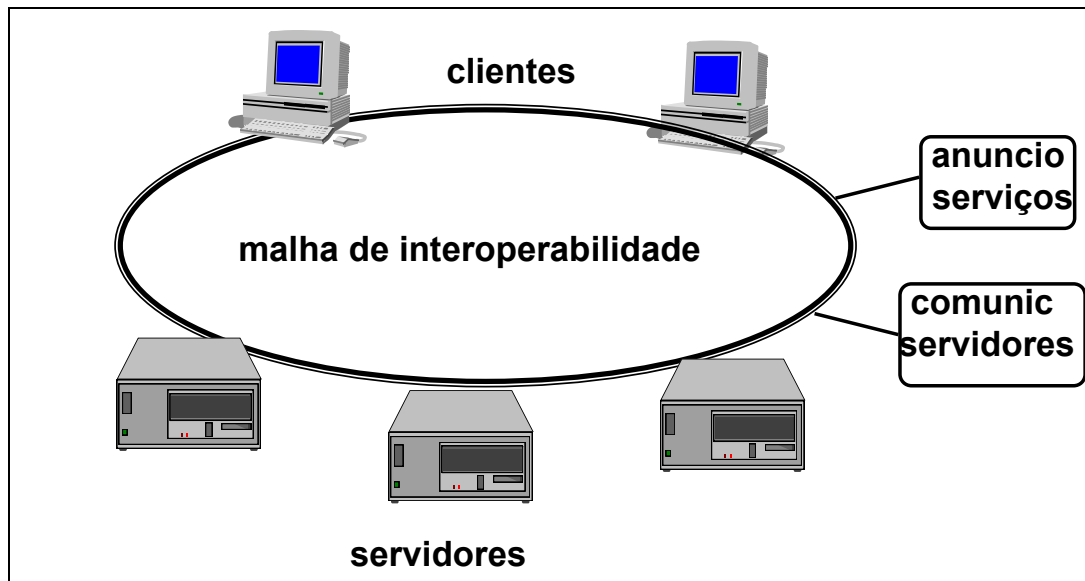


Figura 7.5 - Arquitetura orientada-a-serviços (adaptado de Elkington et al.,1994).

USO DE ARQUITETURAS ORIENTADA A SERVIÇOS PARA SIG

Os princípios de arquitetura orientada-a-serviços podem ser aplicados à arquitetura em camadas descrita na seção 7.3. Os critérios para divisão de serviços entre os clientes e os servidores no ambiente de SIG seguem alguns princípios gerais:

1. Cada servidor será especializado por tipo de representação associada aos objetos geográficos, não se considerando a possibilidade de fragmentar representações.
2. A divisão dos serviços deve levar em conta as enormes diferenças de comportamento e volume de dados entre os diferentes tipos de representação manipuladas por aplicações de Geoprocessamento.
3. Deve-se evitar o tráfego desnecessário de grandes massas de dados na rede. Por exemplo, se desejamos classificar uma imagem LANDSAT de 200 Mb e arquivar o resultado, o serviço de manipulação de imagens deverá estar fisicamente próximo ao serviço de armazenamento de imagens.

Seguindo os princípios acima, propomos a seguir uma possível implementação da arquitetura de referência num ambiente distribuído, seguindo o esquema tradicional de sistemas de gerência de bancos de dados distribuídos heterogêneos, com acoplamento fraco entre os nós (Sheth and Larson, 1990).

Propomos implementar o SAM, SAV e SAC como serviços independentes, especializados por tipo de objeto armazenado (matricial, vetorial ou convencional) em função das enormes diferenças de complexidade e volume entre as diferentes representações. Esta escolha é bastante conveniente quando o volume de dados é substancial. No caso extremo, cada servidor pode ser substituído por um conjunto de servidores, formando um *cluster*, ou ligados por uma rede local ou mesmo por uma rede de longo alcance.

Os servidores especializados (SAM, SAV e SAC) não conhecem a organização lógica do banco de dados geográfico e são responsáveis apenas por serviços elementares de armazenamento e recuperação e pelos serviços locais de gerência de transações, controle de concorrência e controle de integridade.

Os serviços de acesso a representações geométricas (SAM e SAV) devem dispor de estruturas de dados de forma a garantir contigüidade física em memória secundária, condição essencial para se obter um bom desempenho no acesso a dados. Seu projeto deve ser transparente para o usuário, separando organização conceitual de organização física. O SMR será então dividido em serviços de manipulação dedicados a cada tipo de representação (matricial e vetorial).

O serviço de otimização de consultas, e os serviços globais de gerência de transações, controle de concorrência e controle de integridade permanecem de responsabilidade do SMG, implementados (replicadamente) em cada cliente. Em particular, quando uma consulta é submetida a um cliente, a implementação local do SMG deverá estabelecer um plano ótimo de execução e coordenar a sua execução. Cada plano consistirá na execução de subconsultas que resultarão em acessos elementares a geometrias matriciais pelo SAM, geometrias vetoriais pelo SAV e atributos convencionais pelo SAC, repassados a cada um destes pela malha de interoperabilidade. O subsistema de visualização (SV) será implementado por

um serviço separado, em princípio no ambiente do cliente, embora do ponto de vista lógico nada impeça a criação de um servidor especializado em visualização. Assim teremos os seguintes componentes na arquitetura distribuída:

- Um servidor de representações matriciais (imagens e grades), que implementa os serviços de armazenamento de matrizes (SAM) e de manipulação de matrizes (SMM).
- Um servidor de representações vetoriais (polígonos de geometrias complexas), que implementa os serviços de armazenamento de vetores (SAV) e de manipulação de vetores (SMV).
- Um servidor de atributos convencionais (SAC), que implementa os serviços típicos de um SGBD convencional.
- Um serviço de gerência e manipulação de objetos geográficos (SMG), responsável pelo controle de integridade, controle de concorrência e gerência de transações e pela montagem de planos de execução de consultas. Este serviço, por conveniência, será executado em conjunto com o interpretador de LEGAL.
- Serviços executados nos clientes, que incluem um interpretador de LEGAL um serviço de visualização, e uma interface com menus e mouse.

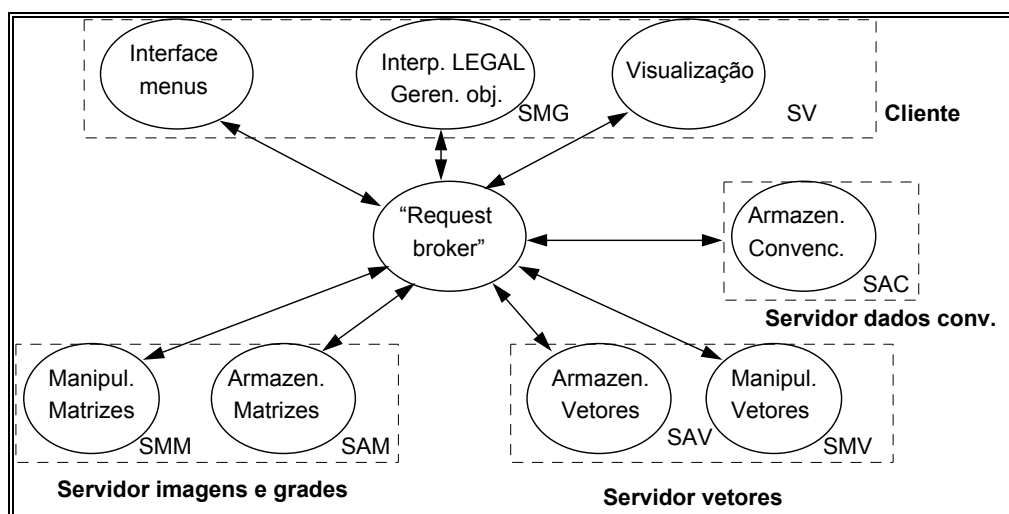


Figura 7.6 - Arquitetura multi-serviços para SIG.

A arquitetura proposta fornece um ambiente aonde cada serviço será executado pelo servidor mais conveniente. As vantagens deste tipo de arquitetura incluem dispor de um gerenciador especializado para cada tipo de aplicação, o que oferece a possibilidade de otimização global do acesso. Dependendo da implementação escolhida, esta arquitetura facilita ainda a ligação com um SGBD relacional convencional.

No próximo capítulo, trataremos da implementação da arquitetura multi-serviços apresentada nesta seção.

8 ARQUITETURA DISTRIBUÍDA PARA SIG: IMPLEMENTAÇÃO

“É interessante contemplar um barranco luxuriante, revestido de muitas plantas diferentes, com pássaros cantando nos galhos, vários insetos esvoaçando e vermes rastejando pela terra úmida, e pensar que estas formas elaboradamente construídas, tão diferentes umas das outras e dependentes entre si de maneira tão complexa, foram todas produzidas por leis que atuam à nossa volta.”

Charles Darwin.

8.1 INTRODUÇÃO

Este capítulo analisa a implementação dos operadores da linguagem LEGAL (definida no capítulo 6) na arquitetura distribuída descrita no capítulo 7, propondo um mapeamento entre as operações da linguagem (próxima do usuário) e as operações atômicas de cada gerenciador. A seção 8.2 discute a estrutura geral dos componentes da arquitetura distribuída e sua implementação. A seção 8.3 apresenta exemplos ilustrativos de operações em LEGAL, mostrando como seriam executadas. A seção 8.4 contém considerações sobre o desempenho da arquitetura proposta, em comparação com as arquiteturas cliente-servidor tradicionalmente utilizadas para implementação de bancos de dados geográficos.

8.2 COMPONENTES DA ARQUITETURA MULTISERVIÇOS

Propomos implementar a arquitetura multiserviços apresentada no capítulo anterior. Para maior conveniência de leitura, repetimos abaixo a estrutura da arquitetura proposta. Nesta seção, iremos discutir a implementação dos diferentes componentes.

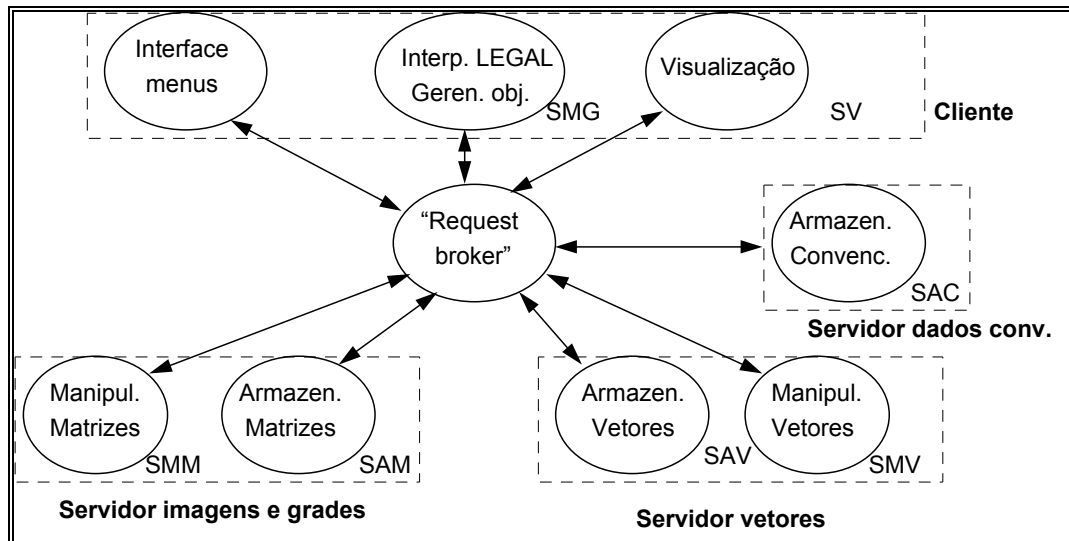


Figura 8.1 - Arquitetura multiserviços.

INTERPRETADOR LEGAL E GERÊNCIA E MANIPULAÇÃO DE OBJETOS GEOGRÁFICOS (SMG)

Este serviço será responsável pelo tratamento dos programas em LEGAL e pela tradução dos comandos da linguagem para o protocolo de comunicação do sistema. O SMG conhece a ligação entre a descrição dos objetos geográficos e sua representação geométrica.

O interpretador opera na máquina cliente, em estreita ligação com o serviço de visualização. A máquina cliente mantém um "buffer" de objetos geográficos, aonde são armazenados (de forma temporária) os objetos geográficos a ser tratados pelo SMG e pelo SV. Os comandos de LEGAL são divididos em três grupos:

- **Recuperação e manipulação de geo-campos:** neste caso, o interpretador não executa o comando, mas o repassa para os serviços SMM, que por sua vez, chama o SAM ou o SAV.
- **Consultas espaciais:** o interpretador estabelece um plano de execução para o comando, e executa subconsultas sobre geometrias matriciais (pelo SAM), sobre geometrias vetoriais (pelo SAV) e sobre atributos convencionais (pelo SAC).

- Apresentação de objetos geográficos: o interpretador repassa as informações sobre os objetos a ser visualizados para o SV; este compõe a apresentação a partir de chamadas aos serviços SAM, SAV e SAC, adicionados de processamentos eventualmente necessários (como agrupar atributos de objetos).

O projeto do interpretador neste ambiente deve seguir princípios semelhantes ao caso clássico de *bancos de dados distribuídos*. Neste caso, considera-se que não há duplicação de informações e que o objeto geográfico está *fragmentado* em seus componentes (atributos e representações geométricas). A gerência das transações, envolvendo mais de um servidor, será realizada através de um protocolo de confirmação em duas fases, ou *two-phase commit* (2PC) (Bernstein et al., 1987).

A versão do protocolo 2PC aplicado ao caso da arquitetura multiservidor para SIG opera como segue:

1. O SMG solicita ao SAC a criação de um novo identificador de objetos geográficos, em seu catálogo de objetos. O SAC responde com um identificador provisório do objeto e um indentificador (também provisório) para as suas representações geométricas.
2. O SMG solicita ao SMM ou SMV (conforme o caso) a realização da operação que cria, a partir do identificador recebido, a representação associada ao objeto geográfico (por exemplo, uma operação de filtragem em imagens).
3. O SMV (ou SMM) sinaliza ao SMG o término com sucesso da operação.
4. O SMG pede então aos serviços envolvidos (SAC e SMM ou SMV) que confirmem a existência do objeto geográfico no seu catálogo.

Se o SMM (ou SMV) sinalizar com erro na execução da operação, o SMG avisa ao SAC para cancelar a operação.

VISUALIZAÇÃO (SV)

Este serviço é executado na máquina cliente, e realiza as funções de processamento da linguagem de apresentação (`SHOW-AS-WHERE-GROUP BY`) e gerenciamento dos eventos de interação, incluindo o atendimento a consultas por apontamento (“quais as características do objeto apontado?”) e a edição gráfica.

O SV é apenas uma *interface de visualização*, pois o acesso às estruturas gráficas (matrizes e vetores) será realizado por cada servidor (SAM e SAV). A comunicação entre o serviço de visualização e estes servidores opera por meio de um protocolo cliente-servidor do tipo “X Window System”: o SV fornece ao servidor apropriado a identificação da janela de desenho, onde será realizada a plotagem da representação associada ao objeto que se deseja visualizar.

Esta estratégia facilita o uso de métodos de acesso multi-resolução. Ao receber um pedido, o SV enviará o comando apropriado ao SMV ou SMM juntamente com o identificador da janela de desenho utilizada, para que estes serviços apresentem os dados nesta janela na resolução compatível. No caso de dados vetoriais, se o mecanismo de acesso utilizado for semelhante ao de uma árvore VR (Mediano et al., 1994) apenas os ramos superiores da árvore serão percorridos, o que pode resultar num ganho significativo de desempenho.

ARMAZENAMENTO E MANIPULAÇÃO VETORIAL (SAV/SMV)

O subsistema de armazenamento vetorial e manipulação vetorial (SAV/SMV) lida com as geometrias do tipo arco-nó-região ou arco-nó. Estas representações, em conformidade com o modelo de dados proposto, estarão sempre associadas a um mapa de geo-objetos, que agrupa polígonos complexos, mantendo as relações de topologia.

Este subsistema realiza operações elementares sobre geometrias vetoriais (ponto-linha-polígono), tais como cômputo de distância entre dois elementos no plano e verificação de propriedades topológicas entre duas representações geométricas (cruzamento, pertinência, adjacência, superposição).

Na implementação do SAV, o aspecto crucial do desempenho será dispor de *métodos de acesso espacial em memória secundária*. Estes métodos serão utilizados nas operações de carga, salvamento e consulta espacial. Para uma visão geral de métodos de acesso espacial, veja-se Cox Junior (1991) e Rezende (1992).

No caso de memória secundária, Mediano et al. (1994) apresentam uma nova proposta de utilizar uma extensão das *árvores-B* para poder apresentar apenas os dados geográficos relevantes para uma determinada escala, sem incorrer em procedimentos desnecessários. Esta estrutura, chamada de *árvore-V*, permite ainda o acesso a dados multiresolução e é muito conveniente como suporte a métodos de navegação de um grande banco de dados.

A indexação espacial deverá dar suporte às operações de seleção espacial e junção espacial. Para o caso aonde as representações vetoriais estão indexadas espacialmente através de *árvores-R*, Brinkhoff et. al. (1993) discutem o percorrimto sincronizado deste tipo de estruturas para a operação de junção espacial.

Este subsistema também implementa um "buffer", onde são armazenados resultados intermediários e auxiliares das consultas espaciais. O "buffer" de geometrias utiliza uma estratégia de "pre-fetching" (Hemerly, 1993b; Cheng and Hurson, 1991): quando um dado é recuperado do banco, faz-se uma estimativa do próximo dado a ser trazido e já se traz também este dado, colocando-o em memória. Se a escolha tiver sido correta, o desempenho do sistema tem significativa melhoria. Áreas contíguas à escolhida seriam trazidas em conjunto para a memória, esperando-se que a próxima requisição seja feita a partir da continuidade geográfica.

ARMAZENAMENTO E MANIPULAÇÃO MATRICIAL (SAM)

Este subsistema é específico para armazenar, recuperar e manipular geometrias na forma matricial ("raster") e será utilizado para tratar imagens e modelos numéricos de terreno na forma de grade. Como observado anteriormente, o subsistema de armazenamento e manipulação matricial

(SMM/SAM) deve ser capaz de realizar processamentos sobre imagens diretamente no servidor sem transferi-las para a memória do cliente.

A literatura tradicional de bancos de dados (p.ex., O.Deux et al., 1992) trata o problema de recuperar dados matriciais de forma superficial, supondo que estes dados estão armazenados em uma estrutura do tipo "bitmap" e com tamanhos reduzidos. No entanto, uma grande gama de funções de processamento de imagens requer acesso contínuo a disco, por ser infactível ter, em memória, imagens tipicamente maiores que 50 Mbytes.

Na prática, a implementação típica do SAM/SMM será feita sobre o sistema de arquivos do sistema operacional, devido ao tamanho e a quantidade destes objetos. Estes arquivos serão relacionados num catálogo armazenado pelo SAC (ver abaixo), mas para maior portabilidade e redundância em caso de falha, cada arquivo matricial é encapsulado numa estrutura auto-descritiva (um exemplo de estrutura auto-descritiva é o formato GRIB, proposto pela Organização Meteorológica Mundial e utilizado nos sistemas SPRING e METVIEW).

O SAM deve dispor de estruturas de indexação espacial e (opcionalmente) de estruturas de acesso a dados em memória terciária. Neste contexto, a noção de um "buffer" de objetos tem uma conotação muito distinta do caso anterior. Numa primeira abordagem, o "buffer" do SMM/SAM irá conter partes de imagens (sub-blocos), que estão sendo utilizadas pelo cliente. Tome-se um exemplo: um processo de segmentação em uma imagem de 100 MBytes (6.000 linhas x 6.000 pontos x 3 bandas) pode operar de maneira mais eficiente se o "buffer" de imagens, em memória, contiver uma sub-imagem (de tamanho 1000 x 1000 x 3 bandas).

Entre os métodos promissores de indexação espacial para imagens, deve-se mencionar o uso de *wavelets* (Chui, 1992) e métodos hierárquicos (Rosenfeld, 1984) que são muito úteis para construir representações multiresolução.

Em sua estrutura interna, o subsistema SMM/SAM será dividido em serviços independentes, que compartilham a mesma estrutura de dados e o "buffer" de

geometrias matriciais. Cada serviço implementa funções de processamento de imagens e grades, como:

- Operações pontuais, locais e zonais sobre mapas temáticos, grades e imagens.
- Segmentação e classificação de imagens.
- Processamento de imagens (filtragem, restauração, transformação IHS-RGB).
- Cálculo de declividade e exposição em grades regulares.

Estas funções não têm conhecimento da semântica do banco de dados geográficos e recebem como argumento apenas o nome dos arquivos envolvidos e do arquivo de saída.

SUBSISTEMA DE ARMAZENAMENTO CONVENCIONAL (SAC)

O SAC não apresenta maiores problemas para sua implementação. Para seu suporte pode-se utilizar tanto um SGBD relacional como um orientado-a-objetos. O SAC utiliza um SGBD para armazenar a lista dos geo-objetos, geo-campos e mapas de objetos presentes, incluindo a indicação dos arquivos associados.

MALHA DE INTEROPERABILIDADE

As decisões mais importantes sobre a implementação do processo que gerencia a malha de interoperabilidade são o modelo de objetos a ser utilizado e o protocolo de comunicação com os demais processos. As alternativas são:

1. Utilizar uma implementação de mercado, onde as opções incluem SOM (Campagnoni, 1994) e OLE (Brockschmidt, 1994).
2. Implementação de uma “casca” adicional sobre as propostas do mercado, de forma a garantir funcionamento com as diversas alternativas.

3. Uso de protocolo dedicado, de rápida implementação, semelhante ao disponível para o sistema METVIEW (Raoult, 1995).

Um dos problemas com os padrões emergentes na área é a incerteza relativa a seu futuro, o que aumenta em muito o risco de qualquer opção¹⁸. De fato, a conhecida revista “Dr. Dobb’s Journal”, em seu número especial sobre o assunto (Valdés, 1994) alerta para a dificuldade de programação nestes ambientes. Adicionalmente, embora especificações como CORBA definam um padrão de interoperabilidade em sistemas distribuídos, algumas suas implementações (como o SOM) operam em um único espaço de endereçamento (Campagnoni, 1994).

Mas talvez o aspecto mais importante é o “vício de origem”: padrões como SOM e OLE têm como objetivo maior o suporte a documentos compostos, típicos de aplicações de automação de escritório, que não tem relação direta com o manuseio de dados geográficos. Não é razoável comparar o encapsulamento de um desenho de 2 Kb, para transferência de um programa de desenho para um editor de texto, com a formatação de uma imagem de 50 Mb, para comunicação entre um aplicativo de processamento de imagens e outro de geração de mapas.

Com base nestas considerações, propomos a implementação de uma malha de interoperabilidade com base no protocolo utilizado no sistema METVIEW (Raoult, 1995). A malha não possui inteligência própria. Ao ser disparada, lê num arquivo de configuração, contendo os serviços disponíveis na rede e os comandos associados a cada serviço. A comunicação entre os processos é completamente assíncrona e todos os processos são clientes e servidores simultaneamente. Cada processo envia um pedido para requisitar ou confirmar a execução de um serviço. A comunicação entre módulos está baseada numa linguagem com a seguinte meta-sintaxe:

¹⁸ Apenas para relembrar, veja-se o que aconteceu com os desenvolvedores de aplicações para janelas em ambientes SUN. Inicialmente, a SUN tinha o ambiente SunView; posteriormente, introduziu o sistema NeWs; a seguir, os programadores foram orientados a desenvolver em Xview, que suportava o padrão OpenWindows. Finalmente, a SUN declarou sua opção “definitiva” pelo padrão Motif. Todos incompatíveis entre si e tudo isto em menos de cinco anos !

```

COMMAND,

    PARAMETER1 = VALUE,

    PARAMETER2 = VALUE1/VALUE2

    PARAMETER3 = VAL1/TO/VAL2/BY/VAL3

```

No caso da arquitetura multiserviços para SIG, indicamos na tabela 8.1 um conjunto mínimo de comandos que compõem o protocolo de comunicação associado à malha de interoperabilidade. Uma descrição completa deste protocolo foge ao escopo deste trabalho, sendo indicados apenas o comando, o serviço que o envia, o serviço de destino e uma descrição de seu significado.

TABELA 8.1
PROTOCOLO DE COMUNICAÇÃO

comando	serviço requisitante	serviço receptor	descrição
BEGIN_TRANSACTION	SMG	malha	solicita início de transação
TRANSACTION_INIT	malha	SMG	inicia uma transação distribuída
END_TRANSACTION	SMG	malha	informa final de transação
CATALOG	SMG	SAC	solicita informações do catálogo
IDENT	SAC	SMG	informações do catálogo
NEW	SMG	SAC	criação de novo objeto
Manipulação Espacial ¹⁹	SMG	SMM	operação de manipulação
COMPLETED	SMM	SMG	término de manipulação
COMMIT	SMG	SAC, SAV, SAM	confirma criação de objeto
COMMITTED	SAC, SAV, SAM	SMG	informa criação definitiva de objeto

¹⁹Na realidade, tratam-se de vários comandos, um para cada operação de manipulação.

ERROR	SAC, SAV, SAM	SMG	informa erro ao realizar parte de uma transação
ABORT	SMG	SAC, SAV, SAM	solicita cancelamento da transação
SELECT	SMG	SAC	consulta por atributos
QUERY_RESULT	SAC	SMG	resultado de consulta por atributos
SPATIAL_OPERATION	SAC	SMV	consulta espacial
SPATIAL_RESULT	SMV	SAC	resultado de consulta espacial
VIEW_INIT	SMG	SV	início de visualização
SHOW	SMG	SV	pedido de visualização
VECTOR_DRAW	SV	SMV	desenho de vetores
RASTER_DRAW	SV	SMM	desenho de matrizes

Por exemplo, um pedido do SMG ao SAC sobre quais as informações associadas a uma imagem de nome “Spot_SJC” seria escrito como:

```
CATALOG,

    TRANSACTION_ID = "951220",

    NAME = "Spot_SJC",

    CLASS = "IMAGE",

    SUBCLASS = "SPOT".
```

Para maiores detalhes sobre esta estratégia de implementação da malha de interoperabilidade no contexto do sistema METVIEW, o leitor deve referir-se a Souza (1994).

8.3 EXEMPLOS DE CONSULTAS EM LEGAL

Para melhor entendimento do funcionamento da arquitetura multiserviços proposta, consideramos três programas em LEGAL:

- cálculo de declividade sobre modelo numérico de terreno;
- seleção espacial sobre geo-objetos com apresentação de atributos de geo-objetos;
- junção espacial com apresentação de atributos de geo-objetos.

Os exemplos têm caráter indicativo. Uma implementação definitiva deste protocolo incluiria aspectos não tratados aqui, como entrada de dados. Consideramos que os parâmetros de cada comando são auto-explicativos no contexto deste documento.

Uma das considerações mais importantes no protocolo é o uso extensivo que faremos dos identificadores dos objetos. Como prevê o modelo de dados apresentado no capítulo 4 e implementado em LEGAL, cada objeto geográfico tem um identificador único. Esta propriedade facilita em muito a passagem de informações entre os serviços. Por exemplo, o comando `QUERY_RESULT` é utilizado pelo SAC para devolver ao SMG uma lista dos objetos que satisfazem a uma consulta não-espacial:

```
QUERY_RESULT,  
  
    TRANSACTION_ID = "960125",  
  
    NUM_ELEMENTS = 3,  
  
    GEOID_LIST = [56748, 56749, 57001].
```


CÁLCULO DE DECLIVIDADE

Considere-se o seguinte programa em LEGAL, que calcula a declividade do Vale do Paraíba, com resolução de 1000 x 1000 metros:

```
VARIABLE topo      Topografia;
VARIABLE decliv    Declividade;

topo:= SELECT gc FROM gc IN Topografia
      WHERE gc.nome = "topoVale92";
decliv:= NEW (NAME = "declVale92")
          REPRESENTED_BY RASTER
          (RES_X=1000m, RES_Y= 1000m);
decliv:= SLOPE (topo);
```

A execução deste programa segue os passos indicados no que segue. Inicialmente, o interpretador solicita à malha de interoperabilidade um identificador de transação (`TRANSACTION_ID`), que será incluído em todos os comandos associados a este programa, e que servirá para que a malha de interoperabilidade mantenha o registro da transação e dirija corretamente a resposta dos servidores (SAC, SAV e SAM) para a máquina-cliente. A malha de interoperabilidade mantém um registro de todas as transações pendentes.

Assim, o SMG faz inicialmente uma requisição de início de transação à malha de interoperabilidade, na forma:

```
BEGIN_TRANSACTION,
      HOST_ID = 150.163.2.4
```

A malha responde então com um comando indicando o identificador da transação:

```
TRANSACTION_INIT,
      TRANSACTION_ID = 951220
```

Ao receber o comando

```
topo:= SELECT gc FROM gc IN Topografia
        WHERE gc.nome = "topoVale92";
```

o SMG requisita informações para o SAC, para determinar o identificador da representação associada. Emite um comando `CATALOG`, com a seguinte sintaxe:

```
CATALOG,
        TRANSACTION_ID = 951220,
        NAME = "topoVale92",
        CLASS = DTM,
        SUBCLASS = "TOPOGRAFIA".
```

Este comando será enviado pela malha de interoperabilidade para o SAC, que procura em seu catálogo o identificador da representação associada, e retorna a informação ao SMG, através de um comando `IDENT`:

```
IDENT,
        TRANSACTION_ID = 951220,
        NAME = "topoVale92",
        NUM_ELEMENTS = 1,
        GEOID_LIST = 99457,
        REPRESENTATION = RASTER,
        NUM_REPRES = 1,
        REPRES_ID_LIST = 99740.
```

O SMG não traz a representação para a memória do cliente, mas prossegue a execução do programa, e ao encontrar o comando

```
decliv:= NEW (NAME = "declVale92")
        REPRESENTED_BY RASTER (RES_X=1000m, RES_Y=
1000m);
```

envia comando para o SAC:

```
NEW,

    TRANSACTION_ID = 951220,

    NAME = "declVale92",

    CLASS = DTM,

    SUBCLASS = "DECLIVIDADE",

    RES_X = 1000m,

    RES_Y = 1000m,

    TRANSACTION_MODE = WAIT.
```

Este comando solicita a criação de uma nova entrada no catálogo de objetos da classe “declividade”. O último parâmetro informa ao SAC que está em curso uma transação longa, e que este deve esperar até o fim da transação (quando será enviado um comando `COMMIT`), para realmente registrar a existência do novo dado. A resposta do SAC é retornar a identificação (provisória) do novo dado ao SMG, através do comando `IDENT`:

```
IDENT,

    TRANSACTION_ID = "951220",

    NUM_ELEMENTS = 1,

    GEOID_LIST = 99732,

    NUM_REPRES = 1,

    REPRESENTATION = RASTER,

    REPRES_ID_LIST = 101597.
```

Ao encontrar o comando em `LEGAL`

```
decliv:= SLOPE (topo);
```

o SMG já possui toda a informação para realizar a operação. Neste momento, envia ao SMM/SAM um comando para a execução do serviço de declividade:

```

SLOPE,

    TRANSACTION_ID = 951220,

    REPRES_ID_IN = 99740,

    REPRES_ID_OUT = 101597,

    RES_X = 1000m,

    RES_Y = 1000m.

```

O SMM/SAM recebe este comando e executa localmente a operação de cálculo de declividade, gerando um arquivo como resultado e retornando um comando ao SMG na forma:

```

COMPLETED,

    TRANSACTION_ID = 951220,

    OPERATION = "SLOPE",

    REPRES_ID_IN = 99740,

    REPRES_ID_OUT = 101597.

```

Ao receber a informação sobre o término da operação, o SMG avisa ao SAC e ao SMM para que registrem o novo geo-campo no seu catálogo:

```

COMMIT,

    TRANSACTION_ID = 951220,

    NUM_OBJECTS = 1,

    GEOID_LIST = 99732,

    NUM_REPRES = 1,

    REPRESENTATION = RASTER,

    REPRES_ID_LIST = 101597.

```

Ao receber este comando, o SAC registra em seu catálogo o novo dado e retorna ao SMG:

```

COMMITTED,

    SERVER = CATALOG,

    TRANSACTION_ID = 951220,

    NUM_ELEMENTS = 1,

    GEOID_LIST = 99732.

```

Por sua vez, o SMM também informa a realização da operação:

```

COMMITTED,

    SERVER = RASTER,

    TRANSACTION_ID = 951220,

    NUM_ELEMENTS = 1,

    GEOID_LIST = 99732,

    NUM_REPRES = 1,

    REPRESENTATION = RASTER,

    REPRES_ID_LIST = 101597.

```

Finalmente, o SMG pode avisar à malha o fim da transação:

```

END_TRANSACTION,

    TRANSACTION_ID = 951220.

```

CONSULTA ESPACIAL SOBRE GEO-OBJETOS

SELEÇÃO ESPACIAL

Consideremos a seguir o seguinte programa em LEGAL, que realiza uma seleção espacial, seguido de apresentação (“recupere todos os postos de saúde com mais de 20 leitos a menos de 15 km do hospital municipal de São José dos Campos e mostre-os com um símbolo apropriado na cor vermelha”).

```

COLLECTION postos    Postos_Saude;
VARIABLE mapa_SJC    Mapa_Viario;

mapa_SJC := ( SELECT mapa FROM MAPA_VIARIO
               WHERE mapa.nome = "Mapa SJCampos");

postos := ( SELECT postos
             FROM postos IN POSTOS_SAUDE ON MAP mapa_SJC,
             hosp IN HOSPITAL ON MAP mapa_SJC,
             WHERE ( hosp.name = "Municipal" AND
                     DISTANCE (postos, hosp ) <  15 km)
             AND postos.leitos > 20);

SET MODE new;

SHOW mapa_SJC AS LINES
      WHERE LINE_COLOUR = "YELLOW";

SET MODE OVERLAY;

SHOW postos.geometry
      AS SYMBOL
      WHERE SYMBOL_NAME = "saude.dxf"
      AND SYMBOL_COLOUR = "red";

```

Como no programa anterior, o SMG faz inicialmente uma requisição de início de transação à malha de interoperabilidade, na forma:

```

BEGIN_TRANSACTION,

      HOST_ID = 150.163.2.4

```

A malha responde então com um comando indicando o identificador da transação:

```

TRANSACTION_INIT,

```

```
TRANSACTION_ID = 960125.
```

Ao receber o comando LEGAL

```
mapa_SJC := ( SELECT mapa FROM mapa IN Mapa_Viario
               WHERE mapa.nome = "Mapa SJCampos");
```

o SMG consulta o catálogo do SAC, para determinar qual o nome do identificador associado ao objeto. Para isto, emite um comando `CATALOG`, com a seguinte sintaxe:

```
CATALOG,

    TRANSACTION_ID = 9601025,

    NAME = "Mapa_Viario_SJC",

    CLASS = CADASTRAL,

    SUBCLASS = "MAPA_VIARIO".
```

Este comando será enviado pela malha de interoperabilidade para o SAC, que procura em seu catálogo o nome do arquivo físico associado, e retorna a informação ao SMG, através de um comando `IDENT`:

```
IDENT,

    TRANSACTION_ID = 960125,

    NAME = "Mapa_Viario_SJC",

    NUM_ELEMENTS = 1,

    GEOID_LIST = [88218],

    REPRESENTATION = VECTOR,

    NUM_REPRES = 1,

    REPRES_ID_LIST = [100027].
```

O SMG prossegue a execução do programa e encontra o comando:

```
postos := ( SELECT postos
FROM postos IN POSTOS_SAUDE ON MAP mapa_SJC,
      hosp IN HOSPITAL ON MAP mapa_SJC,
WHERE(hosp.nome = "Municipal" AND
      DISTANCE(postos, hosp) <15 km
      AND      postos.leitos > 20);
```

Para processar este comando, o SMG precisa inicialmente montar um plano de otimização de consulta e decidir qual a ordem de execução das subconsultas. Neste caso, como o custo de calcular a operação espacial é maior que o custo da seleção por atributos, executa primeiramente a seleção não espacial. Inicialmente, o SMG questiona o SAC para conhecer o identificador do geo-objeto cujo nome é “Hospital Municipal”:

```
CATALOG,
      TRANSACTION_ID = "9601025",
      NAME = "Municipal",
      CLASS = GEOOBJECT,
      SUBCLASS = "HOSPITAL".
```

Ao receber este comando, o SAC verifica seu catálogo e responde:

```
IDENT,
      TRANSACTION_ID = "960125",
      NAME = "Municipal",
      NUM_ELEMENTS = 1,
      GEOID_LIST = [66732].
```

A seguir, o SMG requisita os identificadores dos postos de saúde que satisfazem a cláusula não-espacial:


```

SELECT,

    TRANSACTION_ID = "960125",

    FROM_CLAUSE= "POSTOS_SAUDE",

    WHERE_CLAUSE= "POSTOS_SAUDE.leitos > 10".

```

O SAC executará esta consulta da forma tradicional, fornecendo como resposta a identidade dos postos de saúde que satisfazem o predicado:

```

QUERY_RESULT,

    TRANSACTION_ID = "960125",

    NUM_ELEMENTS = 3,

    GEOID_LIST = [56748, 56749, 57001].

```

Vale lembrar que, em nosso modelo de dados, cada mapa de geo-objetos armazena os identificadores dos geo-objetos nele representados. Assim, as informações obtidas permitem que o SMG solicite ao SMV a realização da operação vetorial (cálculo de distância entre os postos de saúde e o hospital):

```

SPATIAL_OPERATION,

    TRANSACTION_ID = 960125,

    NUM_ELEMENTS_1 = 1,

    GEOID_LIST_1 = [66732],

    OBJECT_MAP_ID_1 = 88218,

    REPRESENTATION_ID_1 = 100027,

    NUM_ELEMENTS_2 = 3,

    GEOID_LIST_2 = [56748, 56749, 57001],

    OBJECT_MAP_ID_2 = 88218,

    REPRESENTATION_ID_2 = 100027,

    OPERATION_TYPE = DISTANCE,

    OPERATION_THRESHOLD = 15 km.

```

O SMV/SAV computará as distâncias entre os diferentes elementos vetoriais, a partir de suas representações. Note-se que as operações espaciais serão sempre calculadas a partir de duas listas de representações vetoriais (cada uma representada em um mapa de objetos). O resultado da operação será um conjunto de pares de objetos que satisfazem o predicado espacial, e será enviado ao SMG:

```
SPATIAL_RESULT,

TRANSACTION_ID = 960125,

NUM_PAIRS = 2,

GEO_ID_LIST = [(66732,56748),(66732,57001)],

DISTANCE_VALUES = [ (5.3 km), (4.5 km) ].
```

No caso de operação de distância, retornamos ainda os valores de distância computados. A partir da lista de pares obtida como resposta da consulta, o SMG compõe a coleção de geo-objetos que satisfazem a consulta. A seguir, passa a analisar o comando em LEGAL:

```
SET MODE new;
```

O resultado deste comando é simplesmente enviar ao SV um comando do tipo:

```
VIEW_INIT,

TRANSACTION_ID = 960125,

VIEW_MODE = NEW.
```

Esta diretiva resultará no apagamento da área de visualização. A seguir tomamos o comando LEGAL:

```
SHOW mapa_SJC AS LINES

WHERE LINE_COLOUR = "YELLOW";
```

Aqui, precisamos recuperar as linhas do mapa cadastral de São José dos Campos, o que requer que o SMG solicite este serviço ao SV:

```

SHOW,

    TRANSACTION_ID = 960125,

    AS_CLAUSE = "LINES",

    WHERE_CLAUSE = "LINE_COLOUR = YELLOW",

    REPRESENTATION = VECTOR,

    CLASS = CADASTRAL,

    NUM_ELEMENTS = 1,

    GEOID_LIST = 88218,

    NUM_REPRES = 1,

    REPRES_ID_LIST = 100027.

```

O SV enviará ao SAV o identificador da janela de visualização associada:

```

VECTOR_DRAW,

    TRANSACTION_ID = 960125,

    X_WINDOW_ID = 465794,

    X_DISPLAY_ID = 150.163.2.4:0,

    X_WINDOW_SIZE_H = 600,

    X_WINDOW_SIZE_Y = 400,

    SCALE = 25000,

    VECTOR_LINE_COLOUR = "YELLOW",

    NUM_REPRES = 1,

    REPRES_ID_LIST = 100027.

```

O SAV/SMV fará o desenho do mapa pedido, na escala apropriada. A seguir, será processado o pedido para apresentar os símbolos relativos aos postos selecionados:

```

SHOW postos.geometry

AS SYMBOL

WHERE SYMBOL_NAME = "saude.dxf"

AND SYMBOL_COLOUR = "red";

```

Ao encontrar este comando LEGAL, o SMG repassa ao SV um comando com a sintaxe:

```

SHOW,

TRANSACTION_ID = 960125,

AS_CLAUSE = "SYMBOL",

WHERE_CLAUSE = "SYMBOL_NAME = saude.dxf",

REPRESENTATION = VECTOR,

CLASS = CADASTRAL,

NUM_ELEMENTS = 2,

GEO_ID_LIST = [ 56743, 57001],

OBJECT_MAP_ID = 88218.

```

O SV entra em contacto com o SAV/SMV para este serviço recupere as representações geométricas associadas aos geo-objetos e as apresente sob a forma de símbolos, com um comando `VECTOR_DRAW`, análogo ao anterior.. A seguir, a transação é encerrada:

```

END_TRANSACTION,

TRANSACTION_ID = 960125.

```

JUNÇÃO ESPACIAL

Considere o seguinte programa em LEGAL (“mostre todos as fazendas improdutivas ocupadas por acampamentos dos sem-terra no estado do São Paulo”).

```
// definicao dos dados
INSTANCE fazenda OF FAZENDA;
INSTANCE acamp OF ACAMPAMENTO;
INSTANCE mapa_fazendas OF MAPA_FAZENDAS;
INSTANCE mapa_acamp OF MAPA_ACAMPAMENTOS;
// definições da coleção
COLLECTION sem_terra OF (ACAMPAMENTO, FAZENDA);
// recuperacao dos mapas
mapa_fazendas:= (SELECT * FROM MAPA_FAZENDAS
                  WHERE NAME = "mapa_faz_SP");
mapa_acamp:= (SELECT * FROM MAPA_ACAMPAMENTOS
              WHERE NAME = "mapa_acampamentos_SP");
// consulta sobre geo-objetos
sem_terra:= ( SELECT acamp, fazenda
              FROM acamp IN ACAMPAMENTO ON MAP mapa_acamp,
                  fazenda IN FAZENDA ON MAP mapa_fazendas
              WHERE      acamp IN fazenda
                  AND fazenda.classe = "improdutiva");
```

Não iremos descrever em detalhe o processamento deste programa, mas fazer algumas considerações sobre os procedimentos de execução da sua parte central (a consulta espacial). Neste caso, o plano de execução também divide a

consulta em dois componentes: as partes convencional e espacial. Da mesma maneira que anteriormente, considera-se que a parte espacial tem custo superior à parte convencional. Deste modo, fazemos inicialmente uma filtragem, selecionando os geo-objetos que satisfazem ao predicado convencional (“fazendas improdutivas”). O resultado será então repassado ao SMV/SAV, para realização da operação de junção espacial.

8.4 CONSIDERAÇÕES SOBRE DESEMPENHO

Comparando-se a arquitetura proposta com as estratégias cliente-servidor descritas no capítulo 7, verificamos que a arquitetura distribuída apresenta a vantagem de ser modular. Cada componente é implementado de forma independente, no ambiente mais adequado para sua operação. Por exemplo, pode-se substituir completamente o SGBD responsável pelo SAC (de um relacional para um orientado-por-objetos), com mínimo impacto no restante do sistema.

Adicionalmente, a arquitetura multiserviços é adaptável, pois pode-se subdividir os gerenciadores especializados (SAM e SAV) em um *cluster* de serviços, envolvendo ainda gerenciamento de dados em memória terciária. Por exemplo, o sistema METVIEW está acoplado a métodos de acesso a memória terciária (Raoult, 1994). Uma discussão adicional sobre o problema de acesso a dados em memória terciária é encontrada em Rosenfeld (1991).

Numa primeira análise, a proposta pode apresentar alguns inconvenientes, principalmente pelo *overhead* associado às transações (pela comunicação entre processos) e pela manutenção da lógica da arquitetura dual (a representação do objeto geográfico continua arquivada de forma separada de seus atributos convencionais). Como mostrado no capítulo anterior, a nova tecnologia de SGBD pós-relacionais e orientados-a-objetos propõe-se justamente a resolver os problemas causados pela arquitetura dual.

Com relação ao *overhead* causado pela comunicação entre processos, consideramos que o preço a pagar é compensado pelo benefício de ter servidores especializados por tipo de dados. Quanto à manutenção da dualidade

representação-atributos em detrimento do uso de novas tecnologias de banco de dados, uma análise feita pelo autor do uso da tecnologia de SGBD orientados-a-objetos em SIG (Câmara, 1994b) levanta dúvidas sobre o desempenho destes ambientes em aplicações ambientais de grande porte.

A arquitetura multiservidor parte da premissa que a dualidade representação-atributos é característica *intrínseca* dos bancos de dados geográficos e que, em lugar de procurar removê-la, deve-se procurar construir a melhor solução que leve em conta as características específicas de cada representação de objetos geográficos.

É forçoso reconhecer que, para bancos de dados geográficos de pequeno e médio porte, arquiteturas cliente-servidor tradicionais (dual ou campos longos) ainda representam uma melhor solução de engenharia que a arquitetura multiservidor. No entanto, o aumento do volume de dados²⁰ inevitável em grandes centros de dados espaciais requer arquiteturas mais poderosas, como a apresentada na proposta.

²⁰ Em final de 1994, o ECMWF já possuía mais de 25 TB de dados matriciais “on-line”, totalizando mais de 5.6 milhões de arquivos independentes (Hoffmann, 1994).

9 COMENTÁRIOS FINAIS

“Quando se trava uma luta prolongada, tenaz e apaixonada começam a delinear-se, geralmente ao fim de certo tempo, os pontos de divergência centrais, essenciais, de cuja solução depende o resultado definitivo da campanha, e em comparação com os quais os episódios menores e insignificantes da luta passam cada vez mais para o segundo plano.”

V. I. Lênin

9.1 CONCLUSÕES

Os capítulos anteriores apresentaram uma possível abordagem para a concepção e desenvolvimento de uma nova geração de sistemas de informação geográfica. Analisando em retrospecto as contribuições apresentadas neste trabalho, alguns pontos nos parecem particularmente relevantes:

- O uso de modelagem orientada-a-objetos permitiu estabelecer uma visão integradora para as duas classes básicas de modelos de informação geográfica (campos e objetos).
- A definição das operações sobre campos e objetos fornece um suporte para a definição de linguagens de consulta e manipulação espacial, análogo ao uso da álgebra relacional para a definição de SQL.
- O padrão OQL (ODMG-93) representa uma base mais adequada para definir uma linguagem de consulta espacial do que a linguagem SQL.
- O uso de arquiteturas distribuídas representa uma alternativa eficaz para a construção de grandes bancos de dados geográficos.

Dentre os temas que podemos apontar, como continuidade das idéias aqui apresentadas, destacam-se:

- A proposição de uma extensão do modelo de dados apresentado no capítulo 4, que incorpore a dimensão temporal.

- A incorporação mais ampla de conceitos de conjuntos nebulosos (“fuzzy sets”) para a definição de campos e objetos geográficos e das operações neles realizadas.
- A definição de operações sobre campos e objetos geográficos, utilizando conceitos de sistemas de bancos de dados dedutivos (Worboys, 1995).
- A definição formal e implementação da linguagem LEGAL, em seus três aspectos (consulta, manipulação e visualização).
- O projeto e a implementação do modelo de dados e de LEGAL, utilizando como suporte um SGBD orientado-a-objetos.
- O projeto e implementação de interfaces orientadas-a-objetos, que utilizem noções de manipulação direta (“drag and drop”) e que reflitam o modelo de dados e as operações de consulta e manipulação propostas.
- O projeto e implementação de arquiteturas distribuídas para bancos de dados geográficos, na linha do apresentado no capítulo 8.

Numa análise dos pontos acima, algumas considerações são pertinentes. Inicialmente, consideramos que a definição de LEGAL está relativamente madura a ponto de merecer esforço para sua implementação (tanto numa arquitetura dual como num SGBD orientado-a-objetos).

Com relação ao conjunto de operações apresentado no capítulo 5, acreditamos que a introdução de conceitos de modelagem temporal, de lógica nebulosa e de sistemas de bancos de dados dedutivos irá certamente engendrar a definição de novas classes de operações sobre dados geográficos.

9.2 À GUIA DE EPÍLOGO

Conforme apontamos anteriormente, este trabalho pretende antes apontar novos caminhos do que esgotar uma tema em particular. Temos a expectativa de que ele possa servir de referência e estímulo para o desenvolvimento de pesquisas e de tecnologia em Geoprocessamento.

Em nossa avaliação, é particularmente relevante o fato do modelo de dados apresentado já estar sendo utilizado como a base conceitual para o desenvolvimento da versão 2.0 do sistema SPRING e de já estar prevista a disponibilidade de LEGAL nas próximas versões deste sistema.

Em conclusão, resta ao autor comentar sua expectativa em relação ao futuro de pesquisas em Geoprocessamento e Bancos de Dados Geográficos. Apesar dos grandes avanços recentes, é forçoso constatar a quantidade de questões em aberto neste campo. Temos o sentimento de que será preciso ainda investir de forma extensiva na análise e formalização dos conceitos de informação espacial, para que nossa disciplina obtenha uma base sólida, que permita passar da atual situação de sistemas desenvolvidos de forma “ad-hoc” para sistemas realmente interoperáveis. Assim, num futuro possível, seriam resolvidos muitos dos problemas práticos hoje associados ao uso e implantação da tecnologia de Geoprocessamento (dependência de fabricantes, perdas semânticas na transferência de dados).

BIBLIOGRAFIA

- Abel, D.J. SIRO-DBMS: A Database Toolkit for Geographical Information Systems. *International Journal on Geographical Information Systems*, 3:103-115, 1989.
- Abel, D.J.; O'Callaghan, J.F. *Directory Facilities for Geographical Data*. CSRIO Internal Report, 1992.
- Alves, D.S. *Modelos de Dados para Sistemas de Informação Geográfica*. Tese de Doutorado. Universidade de São Paulo, 1990.
- Aronoff, S. *Geographical Information Systems: A Management Perspective*. Ottawa, WDI Publications, 1989.
- Assad, E.D.; Sano, E.E. *Sistemas de Informações Geográficas - Aplicações na Agricultura*. Brasília, EMBRAPA, 1993.
- Asrar, G.; Dozier, J. *EOS: Science Strategy for the Earth Observing System*. Woodbury, NY, AIP Press, 1994.
- Bancilhon, F.; Delobel, C.; Kanellakis, P. (eds.) *Building an Object-Oriented Database System: the Story of O2*. San Mateo, CA, Morgan and Kaufmann Publishers, 1991.
- Bentley, J. Multidimensional Search Trees Used for Associative Searching. *Communications of the ACM*, 18:509-517, 1975.
- Bertin, J. *Semiologie Graphique*. Paris, Gauthiers-Villars, 1973.
- Bernstein, P.A.; Hadzilacos, V.; Goodman, N. *Concurrency Control and Recovery in Database Systems*. Reading, Addison-Wesley, 1987.
- Bregt, A. Integrating GIS and Process Models for Global Environmental Assessment. In: International Workshop on Global GIS, Tóquio, Japão, Agosto de 1993. *Proceedings*, Tóquio, ISPRS, pp. 77-84, 1993.
- Brinkhoff, T.; Kriegel, H.P.; Seeger, B. Efficient Processing of Spatial Joins Using R-Trees. In: ACM SIGMOD Conference, Washington, 1993. *Proceedings*, Washington, ACM, pp. 237-246, 1993.
- Brockschmidt, K. OLE Integration Technologies. *Dr. Dobb's Special Report on Interoperable Objects Revolution*, Winter 1994/95, pp. 42-49.
- Burrough, P. A. *Principles of geographic information systems for land resources assessment*. Oxford, Oxford University Press, 1986.

- Burrough, P.A. Development of intelligent geographic information systems. *International Journal of Geographic Information Systems*, 6(1):1-11, 1992.
- Câmara, G.; Velasco, F.R.D.; Oliveira, J.L. Projeto de um Sistema Geográfico de Informação. In: Simpósio Brasileiro de Hardware e Software (SEMISH), Campinas, SP, 1983. *Anais*, Campinas, SBC, 1983.
- Câmara, G.; Souza, R.C.M.; Freitas, U.M.; Casanova, M.A. SPRING: Processamento de Imagens e Dados Georeferenciados. In: V Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, Lindóia, 1992. *Anais*, São José dos Campos, INPE, pp.233-242, 1992.
- Câmara, G.; Souza, R.C.M.; Freitas, U.M.; Paiva, J.A.C. SPRING: Concepção, Evolução, Perspectivas. In: VII Simpósio Brasileiro de Sensoriamento Remoto, Curitiba, PA, 1993. *Anais*, São José dos Campos, INPE, 1993.
- Câmara, G. *Análise de Arquiteturas para Bancos de Dados Geográficos Orientados-a-Objetos*. Monografia de Exame Integrado do Programa de Doutorado em Computação Aplicada do INPE. São José dos Campos, INPE, 1994.
- Câmara, G.; Freitas, U.M.; Cordeiro, J.P.C. Towards an Algebra of Geographical Fields. In: VII Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens. *Anais*, Curitiba, SBC, pp. 205-212, 1994.
- Campagnoni, F.R. IBM's System Object Model. *Dr. Dobb's Special Report on Interoperable Objects Revolution*, Winter 1994/95, pp. 24-28.
- Carey, M.; De Witt, D.J.; Franklin, M.J.; Hall, N.E; McCaulliffe, M.; Naughton, J.F.; Schuh, D.T.; Solomon, M.H.; Tan, C.K.; Tsatalos, O.; White, S.; Zwillig, M.J. Shoring Up Persistent Objects. In: ACM-SIGMOD Conference, 1994. *Proceedings*, ACM, 1994.
- Casanova, M.A.; Hemerly, A.S.; Duarte, M.P.C.; Mediano, M.R.; Câmara, G.; Freitas, U.M. Um Gerenciador para Atributos Espaciais de Objetos Geográficos. In: II Simpósio Brasileiro de Geoprocessamento. *Anais*, USP, São Paulo, SP, 1993.
- Cattell, R.G.G. What are Next-Generation Database Systems?, *Communications of the ACM*, 34(10), October 1991.
- Catell, R.G.G. *The Object Database Standard, ODMG-93*. San Francisco, Morgan Kauffmann, 1994.
- Cheng, J.R.; Hurson, A.R. On the Performance Issues of Object-Based Buffering. In: First International Conference on Parallel and Distributed Information Systems. *Proceedings*, Miami, USA, December 1991.

- Chu, W.W.; Jeong, I.T.; Taira, R.K. A Semantic Modelling Approach for Image Retrieval by Content. *VLDB Journal*, 3:445-447, 1994.
- Chui, C. K. (Ed) *Wavelets: a Tutorial in Theory and Applications*. New York, Academic Press, 1992.
- Clementini, E.; Di Felice, P.; Van Oosterom, P. A Small Set of Formal Topological Relationships Suitable for End-User Interaction. In: Third International Symposium on Spatial Data Handling. *Proceedings*, Singapore, 1993, pp. 277-295.
- Codd, E.F. A Relational Model for Large Shared Data Banks. *Communications of the ACM*, 13(6):377-387, Julho de 1970.
- Cowen, D.J. GIS versus CAD versus DBMS: what are the differences. *Photogrammetric Engineering and Remote Sensing*, 54:1551-4, 1988.
- Cox Jr., F. *Análise de Métodos de Acesso a Dados Espaciais Aplicados a Sistemas Gerenciadores de Banco de Dados*. Dissertação de Mestrado em Computação apresentada ao IMECC-UNICAMP, Campinas, dezembro de 1991.
- David, B.; Raynal, L.; Schorter, G. GeO₂: Why Objects in a Geographical DBMS?. In: *Advances in Spatial Databases (SSD '93)*. Lecture Notes in Computer Science #692, Berlin, Springer-Verlag, 1993.
- David, B.; Voisard, A. *A Unified Approach to Geographic Data Modelling*. Ludwig-Maximilians-Universitat, Institut für Informatik, Munique, Alemanha, Relatório de Pesquisa #9316, Setembro 1993.
- Deux, O. et. al. The O2 System. *Communications of the ACM*, 34(10): 34-48, October 1991.
- Druck, S.; Braga, L.P.V. *Análise e Classificação de Dados de Solos por Geoestatística e Lógica Difusa*. Relatório Interno, Centro Nacional de Pesquisa em Solos, EMBRAPA, 1995.
- vanEck, J.W.; Uffer, M. A Presentation of System 9. *Photogrammetry and Land Information Systems*, 139-178, March 1989.
- Egenhofer, M. *Spatial Query Languages*, Ph.D. Thesis, Depart. of Surveying Engineering, University of Maine, 1989.
- Egenhofer, M. Interaction with GIS via Spatial Queries. *Journal of Visual Languages and Computing*, 1:389-413, 1990.
- Egenhofer, M. Why Not SQL!. *International Journal of Geographic Information Systems*, 6 (2), 1992.

- Egenhofer, M. Spatial SQL: A Query and Presentation Language. *IEEE Transactions on Knowledge and Data Engineering*, 6:86-95, 1994.
- Egenhofer, M. Object Oriented GIS. In: III SIMPOSIO BRASILEIRO DE GEOPROCESSAMENTO. *Anais*. São Paulo, USP, 1995.
- Egenhofer, M.; Clementini, E.; Di Felice, P. Topological Relationships Between Regions with Holes. *International Journal of Geographic Information Systems*, 8(2):129-142, 1994.
- Egenhofer, M.; Frank, A. Object-Oriented Databases: Database Requirements for GIS. In: INTERNATIONAL GIS SYMPOSIUM: THE RESEARCH AGENDA. *Proceedings*, Arlington, VA, Nov. 1987, volume II, p.189-211.
- Egenhofer, M.; Frank, A. Object-Oriented Modeling in GIS: Inheritance and Propagation. In: IX International Symposium on Computer-Assisted Cartography. *Proceedings*, Baltimore, MD, 1989, p. 588-598.
- Egenhofer, M.; Franzosa, R.D. Point-Set Topological Spatial Relations. *International Journal of Geographical Information Systems*, 5 (2): 161-174, 1991.
- Egenhofer, M.; Herring, J. A Mathematical Framework for the Definition of Topological Relationships. In: IV International Symposium on Spatial Data Handling. *Proceedings*, Zurich, 1990, p. 803-813.
- Elkington, M.; Meyer, R.; Mcconaughey, G. *Defining the Architectural Development of EOSDIS to Facilitate Extension to a Wider Data Information System*. Technical Paper, Hugues Applied Information Systems, Landover, Maryland, 1994.
- Ehlers, M., Edwards, G., Berard, I. Integration of Remote Sensing with Geographic Information Systems: A Necessary Evolution. *Photogrammetry Engineering and Remote Sensing*, 55 (11): 1619-1627, 1989.
- Erthal, G.; Alves, D.S.; Câmara, G. Modelo de Dados Geo-Relacional: Uma Visão Conceitual de um Sistema Geográfico de Informações. In: I Simpósio Brasileiro de Computação Gráfica e Processamento De Imagens. *Anais*, Rio de Janeiro, COPPE/UFRJ, 1988.
- ESRI, *ARC/INFO 6.0 Command Reference: GRID*. Redlands, California, USA, 1991.
- Feutchwanger, M. *Towards a Geographic Semantic Database Model*. Ph.D. Thesis, Simon Fraser University, Canada, 1993.

- Frank, A.; Mark, D. Language Issues for GIS. In: Maguire, D.; Goodchild, M.; Rhind, D. (eds) *Geographical Information Systems: Principles and Applications*. New York, John Wiley and Sons, 1991, cap. 11, p. 147-163.
- Freeman, J. The Modelling of Spatial Relations. *Computer Graphics and Image Processing*, 4: 156-171, 1975.
- Goodchild, M. A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems*, 1 (4): 327-334, 1987.
- Goodchild, M. Geographical information science. *International Journal of Geographical Information Systems*, 6 (2): 35-45, 1992a.
- Goodchild, M. Geographical data modeling. *Computers & Geosciences*, 1992, 18(4): 401-408, 1992b.
- Goodchild, M., Haining, R.P., Wise, S. Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*, 6 (4): 407-423, 1992.
- Goodchild, M.; Parks, B.; Steyart, L. *Environmental Modelling with GIS*, Oxford, Oxford University Press, 1993.
- Gomes, J.M.; Velho, L. *Computação Visual: Imagens*. Rio, SBM, 1995.
- Gregory, K.J. *A Natureza da Geografia Física*. Rio de Janeiro, Bertrand Brasil, 1992.
- Günther, O.; Riekert, W.-F. The Design of GODOT: An Object-Oriented Geographic Information System. *Data Engineering Bulletin*, IEEE Computer Society, September 1993.
- Güting, R.H. An Introduction to Spatial Database Systems. *VLDB Journal*, 3(4), October 1994.
- Güting, R.H.; Schneider, M. Realms: a Foundation for Spatial Data Types in Database systems. In: *Advances in Spatial Databases (SSD'93)*. Lecture Notes in Computer Science #692, Berlin, Springer-Verlag, 1993.
- Güting, R.H.; Ridder, T.; Schneider, M. Implementation of the ROSE Algebra: Efficient Algorithms for Realm-based Spatial Data Types. In: *Advances in Spatial Databases (SSD'95)*. Lecture Notes in Computer Science #951, pp. 216-239. Berlin, Springer-Verlag, 1995.
- Gutman, A. R-trees: a dynamic index structure for spatial searching. In: ACM SIGMOD CONFERENCE. *Proceedings*, Boston, ACM, pp. 47-57, 1984.

- Hass, L.; Cody, W. Exploiting extensible DBMS in integrated geographic information systems. In: *Advances in Spatial Databases (SSD'91)*. Lecture Notes in Computer Science #525. Berlin, Springer-Verlag, 1991.
- Hemerly, A.S. *Um Estudo de Linguagens de Consulta Espaciais*. Memorando interno, Centro Científico Rio, IBM Brasil, 1993a.
- Hemerly, A.S. *Um Estudo sobre a Gerência de Objetos Geográficos*. Memorando interno, Centro Científico Rio, IBM Brasil, 1993b.
- Hemerly, A.S.; Mediano, M.; Souza, P.; Tocantins, C.; Duarte, M. *Uma Arquitetura para um Gerenciador de Objetos Geográficos*. Documento Interno, Centro Científico Rio, IBM Brasil, 1994.
- Hoffmann, G. ECMWF Computer Division Status and Plans. *ECMWF Newsletter*, 68:37-40, Winter 1994.
- INTERGRAPH, MGE- *The Modular GIS Environment*, 1990.
- Joly, P. *A Cartografia*. Campinas, Papirus, 1990.
- Kemp, K. *Environmental Modelling with GIS: A Strategy for Dealing with Spatial Continuity*. Ph.D. thesis, Department of Geography, University of Santa Barbara, 1993. NCGIA Technical Report 93-3.
- Lécluse, C.; Richard, P.; Vélez, F. O₂, an Object-Oriented Data Model. In: Bancilhon, F.; Delobel, C.; Kanellakis, P. (eds.) *Building an Object-Oriented Database System: the Story of O2*. San Mateo, CA, Morgan and Kaufmann Publishers, 1991, cap.4, pp.77-95.
- Maguire, D. An Overview and Definition of GIS. In: Maguire, D.; Goodchild, M.; Rhind, D. (eds) *Geographical Information Systems: Principles and Applications*. New York, John Wiley and Sons, 1991, pp. 9-20.
- Maguire, D.; Dangerramond, J. The Functionality of GIS. In: Maguire, D.; Goodchild, M.; Rhind, D. (eds) *Geographical Information Systems: Principles and Applications*. New York, John Wiley and Sons, 1991, pp. 319-35.
- Maguire, D.; Goodchild, M.; Rhind, D. (eds.) *Geographical Information Systems: Principles and Applications*. New York, John Wiley and Sons, 1991.
- Mediano, M. *V-Trees: Um Método de Armazenamento para Dados Vetoriais Longos*. Dissertação de Mestrado, Departamento de Informática, PUC-RJ, 1995.
- Mediano, M.; Casanova, M.; Dreux, M. V-Trees - A Storage Method for Long Vector Data. In: International Symposium on Very Large Data Bases, Chile, 1994. *Proceedings of VLDB '94*, Santiago, 1994.

- Milne, P. ; Milton, S.; Smith, J.L. Geographical Object-Oriented Data Bases: a Case Study. *International Journal of Geographic Information Systems*, 7(1):39-55, 1993.
- Monmonier, M. *Mapping It Out*. Chicago, Chicago University Press, 1992.
- Morehouse, S. The ARC/INFO Geographical Information System. *Computers & Geosciences*, 18(4): 435-443, 1992.
- NCGIA, The Research Plan for the NCGIA, *International Journal of Geographic Information Systems*, 3(2):117-136, 1989.
- Ooi, B.C. *Efficient Query Processing in Geographic Information Systems*. Berlim, Springer-Verlag, 1990. (Lecture Notes on Computer Science, 471).
- O₂ Technology. *OQL (Object Query Language) Manual*. Paris, O₂ Technology, 1995.
- Openshaw, S. An automated geographical analysis system. *Environment and Planning*, 19:431-436, 1987.
- Paredes, R. *Introdução aos Sistemas de Informação Geográfica*. São Paulo, Érica, 1995.
- Pissinou, N.; Makki, K.; Park, E. Towards the design and development of a new architecture for geographic information systems. In: 2nd. International Conference on Information and Knowledge Management - CIKM. *Proceedings*, pp. 537-546, 1993.
- Puppo, E.; Dettori, G. Towards a Formal Model for Multiresolution Spatial Maps. In: *Advances in Spatial Databases (SSD'95)*. Lecture Notes in Computer Science #951, pp. 152-169. Berlin, Springer-Verlag, 1995.
- Raoult, B. Workstation MARS: Access to ECMWF's Data Base. IV Workshop on Meteorological Operational Systems. *Proceedings*, ECWMF, 1994.
- Raoult,B.; Norris,B.; Daabeck, J.; Souza, R.C.M.; Câmara, G. Distributed Architectures for Environmental Visualisation Systems. In: VII Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens. *Anais*, SBC, pp. 249-256, 1995.
- Rezende, M.A.M. *Métodos de Acesso para Bancos de Dados Espaciais*. Relatório Técnico CCR-142, Centro Científico Rio, IBM Brasil, Janeiro 1992.
- Rigaux, P.; Scholl,M. Multi-Scale Partitions: Application to Spatial and Statistical Databases. In: *Advances in Spatial Databases (SSD '95)*. Lecture Notes in Computer Science #951, pp. 170-183. Berlin, Springer-Verlag, 1995.

- Rosenfeld, A. *Multiresolution Image Processing and Analysis*. New York, Springer-Verlag, 1984.
- Samet, H. *The Design and Analysis of Spatial Data Structures*. Reading, Addison-Wesley, 1990.
- Scholl, M.; Voisard, A. Object-oriented databases systems for geographic applications: an experiment with O₂. In: Bancilhon, F.; Delobel, C.; Kanellakis, P. (eds.) *Building an Object-Oriented Database System: the Story of O₂*. San Mateo, CA, Morgan and Kaufmann Publishers, 1991, cap. 28, pp. 585-625.
- Sheth, A.; Larson, J. Federated Database Systems for Managing Distributed, Heterogeneous and Autonomous Databases. *ACM Computing Surveys*, 22(3): 183-236, September 1990.
- Smith, T.R.; Menon, S.; Star, J.; Estes, J.E. Requirements and Principles for the Implementation and Construction of Large-Scale Geographical Information Systems. *International Journal of Geographical Information Systems*, 1:13-31, 1987.
- Smith, T.R.; Frank, A. *Report on Workshop on Very Large Spatial Databases*. NCGIA Technical Paper 89-13, Santa Barbara, CA, 1989.
- Sondheim, M. Modelling the Real World. In: GIS '93 SYMPOSIUM. *Proceedings*, Vancouver, CA, pp. 1-21, 1993.
- Souza, R.C.M.; Câmara, G.; Alves, D.S. O Desenvolvimento de Sistemas de Informação Geográfica e de Processamento Digital de Imagens no INPE. In: I Simpósio Brasileiro de Geoprocessamento, SÃO PAULO, 1990. *Anais*, EPUSP, São Paulo, 1990, pp. 168-173.
- Souza, R.C.M. *METVIEW Internal Architecture*. ECMWF Internal Document. Reading, ECWMF, 1994.
- Star, J.; Estes, J. *Geographic Information Systems: An Introduction*. New York, Prentice-Hall, 1990.
- Stonebraker, M. Managing Persistent Objects in a Multi-Level Store. In: ACM SIGMOD Conference, 1991. *Proceedings*, ACM, 1991.
- Stonebraker, M.; Dozier, J. *SEQUOIA 2000: Large Capacity Object Servers to Support Global Change Research*. Report #91/1, Sequoia 2000 Project, University of California, Berkeley, 1992.
- Stonebraker, M.; Kemnitz, G. The PostGRES Next Generation Database Management System. *Communications of the ACM*, 34(10), October 1991.

- Surveys and Resource Mapping Branch, *Spatial Archive and Interchange Format (SAIF): Profile #1*. Ministry of Environment, Lands and Parks, Province of British Columbia, 1993.
- Tateishi, R. (ed.) *International Workshop on Global GIS*. Proceedings. Tokio, ISPRS Working Group IV/6, 1993.
- Times, V.C; Salgado, A.C. Uma Modelagem Orientada a Objetos para Aplicações Geográficas. In: Simpósio Brasileiro de Banco de Dados, 1995. *Anais*, SBC, 1995.
- Tocantins, C.; Casanova, M.A. Storage and Processing of Vectorial Maps on Parallel Architectures. In: Simpósio Brasileiro de Banco de Dados, 1995. *Anais*, SBC, 1995.
- Tomlin, D. *Geographic information systems and Cartographic Modeling*. Prentice Hall, New York, 1990.
- Tufte, E.R.. *The Visual Display of Quantitative Information*. Cheshire, USA, Graphics Press, 1983.
- Váldez, R. Implementing Interoperable Objects. *Dr. Dobb's Special Report on Interoperable Objects Revolution*, Winter 1994/95, pp. 62-72.
- Vijbrief, T.; vanOosterom, P. The GEO++ System: An Extensible GIS. In: International Symposium on Spatial Data Handling, Arlington, VA, 1992. *Proceedings*, ACM, 1992.
- deWitt, D.; Luo, J.; Patel, J.; Yu, J.-B. Paradise - A Parallel Geographic Information System. In: ACM Workshop on Advances in Geographical Information Systems. Arlington, Virginia, 1993. *Proceedings*, ACM, 1993.
- deWitt, D.; Kabra, N.; Luo, J.; Patel, J.; Yu, J.-B. Client-Server Paradise. In: 20th Vldb Conference, Santiago, Chile, 1994. *Proceedings*, 1994.
- Worboys, M. Object-oriented approaches to geo-referenced information. *International Journal of Geographic Information Systems*, 8(4):385-400, July-August 1994.
- Worboys, M.F.; Hearnshaw, H.M.; Maguire, D.J. Object-oriented data modelling for spatial data bases. *International Journal of Geographical Information Systems*, 4:369-383, 1990.

APÊNDICE

ANÁLISE DOS RELACIONAMENTOS TOPOLÓGICOS

DESCRIÇÃO

Este apêndice contém a análise exaustiva aos relacionamentos topológicos descritos no texto principal do Capítulo 5. Iremos discutir os casos área-área, linha-linha, linha-área, ponto-área, ponto-linha e ponto-ponto. Nossas ferramentas de análise serão a matriz de 4-intersecções e a análise da intersecção dos conjuntos de pontos.

O que faz os relacionamentos topológicos diferir da teoria de conjuntos tradicional é o uso das noções de fronteira e interior.

RELACIONAMENTOS ÁREA-ÁREA

A tabela A.1 indica os 16 possíveis casos de relacionamento área-área, ilustrados na figura A.1. Estes relacionamentos foram agrupados em quatro conceitos: *disjoint*, *overlap*, *touch* e *in*. Por simplicidade, não utilizamos (como faz Egenhofer) os conceitos de “equal”, “covers” e “covered by”.

1. *DISJOINT* - duas áreas são disjuntas quando não tem pontos em comum.

$$A \text{ disjoint } B \Leftrightarrow A \cap B = \emptyset.$$

2. *IN* - uma área está contida em outra quando a intersecção dos dois conjuntos de pontos é a própria área.

$$A \text{ in } B \Leftrightarrow (A \cap B = A).$$

3. *OVERLAP* - duas áreas se sobrepõem quando seus interiores se tocam e a intersecção dos dois conjuntos não é nenhuma das duas áreas.

$$A \text{ overlap } B \Leftrightarrow (A^\circ \cap B^\circ \neq \emptyset) \wedge (A \cap B \neq A) \wedge (A \cap B \neq B).$$

4. *TOUCH* - duas áreas se tocam quando apenas suas fronteiras se interceptam.

$$A \text{ touch } B \Leftrightarrow (A \cap B \neq \emptyset) \wedge (A^\circ \cap B^\circ = \emptyset).$$

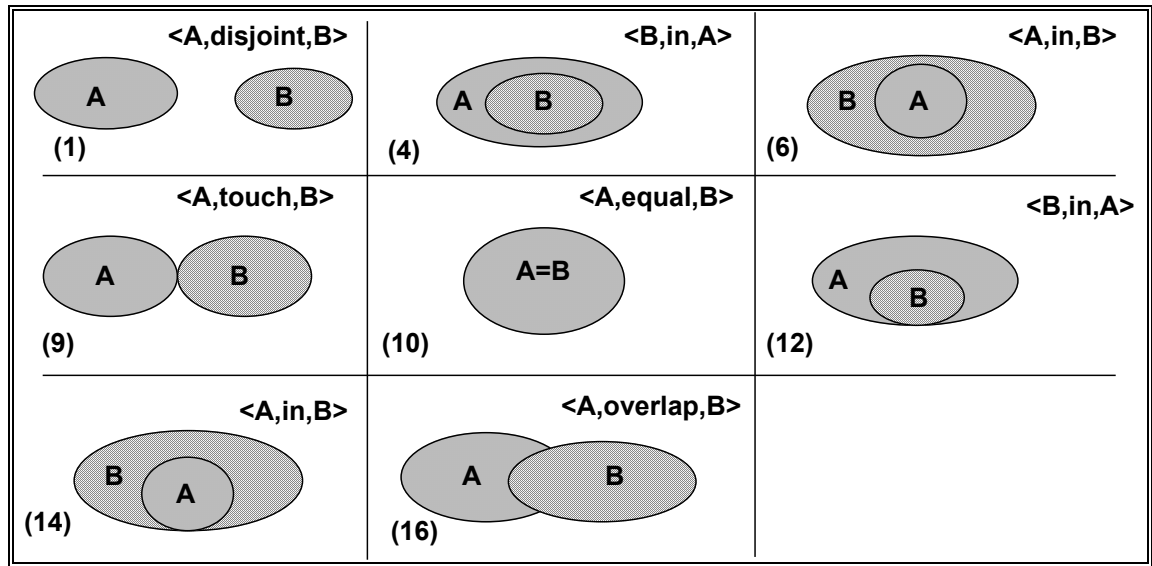


Figura A.1 - Relacionamentos área-área

TABELA A.1 - Relacionamentos área-área

caso	$\delta A \cap \delta B$	$\delta A \cap B^o$	$A^o \cap \delta B$	$A^o \cap B^o$	$A \cap B$	visão intuitiva
1	\emptyset	\emptyset	\emptyset	\emptyset	-	A disjoint B
2	\emptyset	\emptyset	\emptyset	$\neg \emptyset$	-	-
3	\emptyset	\emptyset	$\neg \emptyset$	\emptyset	-	-
4	\emptyset	\emptyset	$\neg \emptyset$	$\neg \emptyset$	B	B in A
5	\emptyset	$\neg \emptyset$	\emptyset	\emptyset	-	-
6	\emptyset	$\neg \emptyset$	\emptyset	$\neg \emptyset$	A	A in B
7	\emptyset	$\neg \emptyset$	$\neg \emptyset$	\emptyset	-	-
8	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	-	-
9	$\neg \emptyset$	\emptyset	\emptyset	\emptyset	-	A touch B
10	$\neg \emptyset$	\emptyset	\emptyset	$\neg \emptyset$	A, B	A in B \wedge B in A
11	$\neg \emptyset$	\emptyset	$\neg \emptyset$	\emptyset	-	-
12	$\neg \emptyset$	\emptyset	$\neg \emptyset$	$\neg \emptyset$	B	B in A
13	$\neg \emptyset$	$\neg \emptyset$	\emptyset	\emptyset	-	-
14	$\neg \emptyset$	$\neg \emptyset$	\emptyset	$\neg \emptyset$	A	A in B
15	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	\emptyset	-	-
16	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neq A, \neq B$	A overlap B

RELACIONAMENTOS LINHA-ÁREA

A tabela A.2 abaixo indica os 14 relacionamentos topológicos possíveis (segundo a matriz de 4-intersecções) entre uma linha e uma área. Estes relacionamentos foram condensados nos conceitos *disjoint*, *touch*, *cross* e *in*. Ocorrem três situações de ambiguidade entre os relacionamentos *cross* e *in*, resolvidas por intermédio do cálculo da intersecção entre os conjuntos de pontos. Estes relacionamentos estão ilustrados na figura A.2.

1. $L \text{ disjoint } A$ - uma linha e uma área são disjuntas quando não tem pontos em comum. $L \text{ disjoint } A \Leftrightarrow L \cap A = \emptyset$.

2. $L \text{ in } A$ - uma linha está contida em uma área quando seus interiores se tocam e a intersecção dos dois conjuntos de pontos é a própria linha.

$$L \text{ in } A \Leftrightarrow (L^\circ \cap A^\circ = \neg \emptyset) \wedge (L \cap A = L).$$

3. $L \text{ cross } A$ - uma linha cruza uma área quando seus interiores se tocam e a intersecção dos dois conjuntos não é a própria linha (isto é, a linha está parcialmente fora da área) $L \text{ cross } A \Leftrightarrow (L^\circ \cap A^\circ = \neg \emptyset) \wedge (L \cap A \neq L)$.

4. $L \text{ touch } A$ - uma linha toca uma área quando a fronteira da área intercepta a fronteira (ou o interior) da linha e os interiores não se tocam. $L, \text{touch}, A \Leftrightarrow \{(\delta L \cap \delta A = \neg \emptyset) \vee (L^\circ \cap \delta A = \neg \emptyset)\} \wedge (L^\circ \cap A^\circ = \emptyset)$.

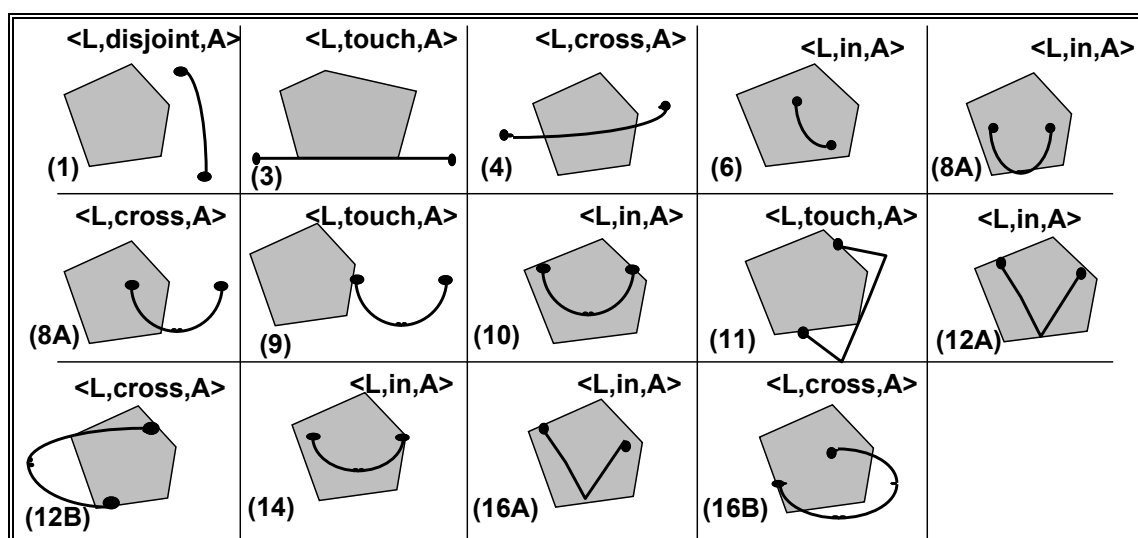


Figura A.2 - Relacionamentos possíveis linha-área

TABELA A.2 - RELACIONAMENTOS LINHA-ÁREA

caso	$\delta L \cap \delta A$	$\delta L \cap A^\circ$	$L^\circ \cap \delta A$	$L^\circ \cap A^\circ$	$L \cap A$	relacionamen to
1	\emptyset	\emptyset	\emptyset	\emptyset	-	L disjoint A
2	\emptyset	\emptyset	\emptyset	$\neg \emptyset$	-	-
3	\emptyset	\emptyset	$\neg \emptyset$	\emptyset	-	L touch A
4	\emptyset	\emptyset	$\neg \emptyset$	$\neg \emptyset$	-	L cross A
5	\emptyset	$\neg \emptyset$	\emptyset	\emptyset	-	-
6	\emptyset	$\neg \emptyset$	\emptyset	$\neg \emptyset$	-	L in A
7	\emptyset	$\neg \emptyset$	$\neg \emptyset$	\emptyset	-	-
8A	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	L	L in A
8B	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neq L$	L cross A
9	$\neg \emptyset$	\emptyset	\emptyset	\emptyset	-	L touch A
10	$\neg \emptyset$	\emptyset	\emptyset	$\neg \emptyset$	-	L in A
11	$\neg \emptyset$	\emptyset	$\neg \emptyset$	\emptyset	-	L touch A
12A	$\neg \emptyset$	\emptyset	$\neg \emptyset$	$\neg \emptyset$	L	L in A
12B	$\neg \emptyset$	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neq L$	L cross A
13	$\neg \emptyset$	$\neg \emptyset$	\emptyset	\emptyset	-	-
14	$\neg \emptyset$	$\neg \emptyset$	\emptyset	$\neg \emptyset$	-	L in A
15	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	\emptyset	-	-
16A	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	L	L in A
16B	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neq L$	L cross A

RELACIONAMENTOS LINHA-LINHA

A tabela A.3 abaixo indica os 24 relacionamentos topológicos possíveis (segundo a matriz de 4-intersecções) entre duas linhas, ilustrados na figura A.3. Para caracterizar melhor os relacionamentos e ambiguidades será preciso lançar mão da noção de dimensão da intersecção. Note-se que no caso de relações linha-linha, quatro tipos de relações são possíveis:

1. L_1 disjoint L_2 - duas linhas são disjuntas quando não tem pontos em comum. L_1 disjoint $L_2 \Leftrightarrow L_1 \cap L_2 = \emptyset$.

2. L_1 *cross* L_2 - duas linhas se cruzam quando seus interiores se tocam e sua intersecção define um conjunto de pontos de dimensão zero (isto é, um conjunto de pontos esparsos).

$$L_1 \text{ cross } L_2 \Leftrightarrow L_1^\circ \cap L_2^\circ = \neg \emptyset \wedge \dim(L_1 \cap L_2) = 0.$$

3. L_2 *overlap* L_1 - duas linhas se sobrepõem quando seus interiores se tocam e sua intersecção define um conjunto de pontos de dimensão um (isto é, sua intersecção contém pelo menos uma linha).

$$L_2 \text{ overlap } L_1 \Leftrightarrow (L_1^\circ \cap L_2^\circ = \neg \emptyset) \wedge (\dim(L_1 \cap L_2) = 1).$$

4. L_1 *touch* L_2 - duas linhas se tocam quando a fronteira de uma linha intercepta a fronteira (ou o interior) da outra e os interiores não se tocam.

$$L_1 \text{ touch } L_2 \Leftrightarrow (L_1 \cap L_2 = \neg \emptyset) \wedge (L_1^\circ \cap L_2^\circ = \emptyset).$$

Note-se que incluímos na relação “overlap” os casos aonde uma linha está completamente contida em outra. Isto foi feito por tanto por tornar mais clara a formalização, como por corresponder a uma noção cognitiva; a noção de “inclusão” é mais evidente quando lidamos com dois conjuntos de pontos de dimensões distintas (e.g, linha em área).

TABELA A.3 - Relacionamentos linha-linha

<i>caso</i>	$\delta L_1 \cap \delta L_2$	$\delta L_1 \cap L_2^\circ$	$L_1^\circ \cap \delta L_2$	$L_1^\circ \cap L_2^\circ$	$\dim(L_1 \cap L_2)$	<i>relacionamento</i>
1	\emptyset	\emptyset	\emptyset	\emptyset	-	L_1 disjoint L_2
2A	\emptyset	\emptyset	\emptyset	$\neg \emptyset$	1	L_1 overlap L_2
2B	\emptyset	\emptyset	\emptyset	$\neg \emptyset$	0	L_1 cross L_2
3	\emptyset	\emptyset	$\neg \emptyset$	\emptyset	-	L_2 touch L_1
4A	\emptyset	\emptyset	$\neg \emptyset$	$\neg \emptyset$	1	L_2 overlap L_1
4B	\emptyset	\emptyset	$\neg \emptyset$	$\neg \emptyset$	0	L_1 cross L_2
5	\emptyset	$\neg \emptyset$	\emptyset	\emptyset	-	L_1 touch L_2
6A	\emptyset	$\neg \emptyset$	\emptyset	$\neg \emptyset$	1	L_1 overlap L_2
6B	\emptyset	$\neg \emptyset$	\emptyset	$\neg \emptyset$	0	$\langle L_1, \text{cross}, L_2 \rangle$
7	\emptyset	$\neg \emptyset$	$\neg \emptyset$	\emptyset	-	$\langle L_1, \text{touch}, L_2 \rangle$
8A	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	1	$\langle L_1, \text{overlap}, L_2 \rangle$
8B	\emptyset	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	0	$\langle L_1, \text{cross}, L_2 \rangle$
9	$\neg \emptyset$	\emptyset	\emptyset	\emptyset	-	$\langle L_1, \text{touch}, L_2 \rangle$
10A	$\neg \emptyset$	\emptyset	\emptyset	$\neg \emptyset$	1	$\langle L_1, \text{overlap}, L_2 \rangle$
10B	$\neg \emptyset$	\emptyset	\emptyset	$\neg \emptyset$	0	L_1 cross L_2
11	$\neg \emptyset$	\emptyset	$\neg \emptyset$	\emptyset		L_1 touch L_2
12A	$\neg \emptyset$	\emptyset	$\neg \emptyset$	$\neg \emptyset$	1	L_1 overlap L_2
12B	$\neg \emptyset$	\emptyset	$\neg \emptyset$	$\neg \emptyset$	0	L_1 cross L_2
13	$\neg \emptyset$	$\neg \emptyset$	\emptyset	\emptyset	-	L_1 touch L_2
14A	$\neg \emptyset$	$\neg \emptyset$	\emptyset	$\neg \emptyset$	1	$L_1, \text{overlap} L_2$
14B	$\neg \emptyset$	$\neg \emptyset$	\emptyset	$\neg \emptyset$	0	L_1 cross L_2
15	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	\emptyset		L_1 touch, L_2
16A	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	1	L_1 overlap L_2
16B	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	$\neg \emptyset$	0	L_1 cross L_2

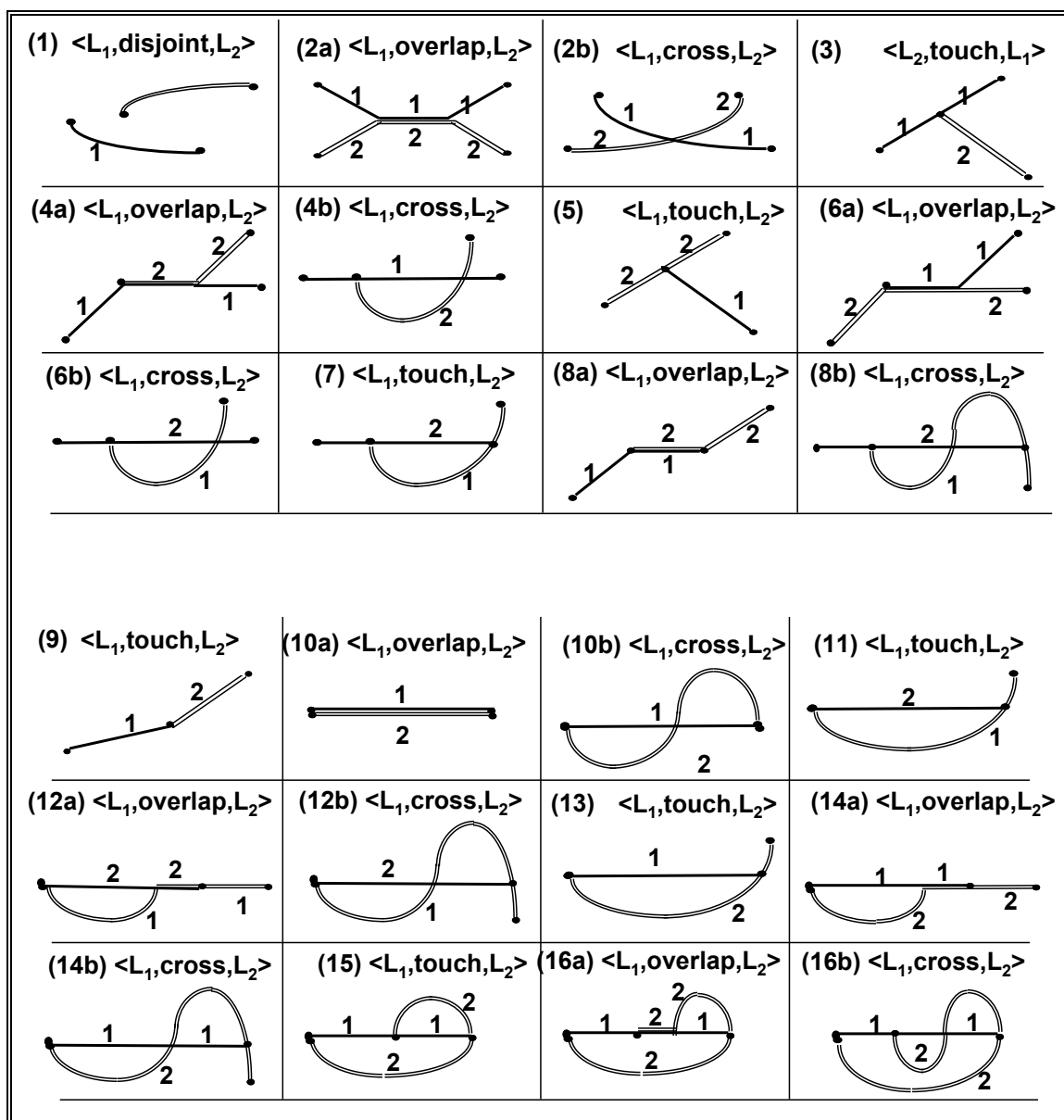


Figura A.3 - Relacionamentos linha-linha.

RELACIONAMENTOS PONTO-ÁREA

No caso de relacionamentos ponto-área, o número de casos a verificar é bastante restrito, pois a fronteira de um ponto é vazia. A tabela A.4 traz os 3 casos possíveis, ilustrados na figura A.4.

1. *P disjoint A* - um ponto e uma área são disjuntos quando sua intersecção é vazia.

$$P \text{ disjoint } A \Leftrightarrow P \cap A = \emptyset.$$

2. *P in A* - um ponto está dentro de uma área quando o ponto e o interior da área tem intersecção.

$$P \text{ in } A \Leftrightarrow P^\circ \cap A^\circ = \neg \emptyset.$$

3. *P touch A* - um ponto toca uma área quando o ponto intercepta a fronteira da área.

$$P \text{ touch } A \Leftrightarrow P^\circ \cap \delta A = \neg \emptyset.$$

Tabela A.4 - Relacionamentos ponto-área

<i>cas o</i>	$P^\circ \cap \delta A$	$P^\circ \cap A^\circ$	<i>relacionamento</i>
1	\emptyset	\emptyset	<i>P disjoint A</i>
2	\emptyset	$\neg \emptyset$	<i>P inside A</i>
3	$\neg \emptyset$	\emptyset	<i>P touch A</i>
4	$\neg \emptyset$	$\neg \emptyset$	-

9.3 RELACIONAMENTOS PONTO-LINHA

Também neste caso, o universo a considerar é restrito, como mostra a tabela A.5.

1. *P disjoint L* - um ponto e uma linha são disjuntos quando sua intersecção é vazia.

$$P \text{ disjoint } L \Leftrightarrow P \cap L = \emptyset.$$

2. *P in L* - um ponto está dentro de uma linha quando o ponto e o interior da linha tem intersecção.

$$P \text{ in } L \Leftrightarrow P^\circ \cap L^\circ = \neg \emptyset.$$

3. *P touch L* - um ponto toca uma linha quando o ponto intercepta a fronteira da linha.

$$P \text{ touch } L \Leftrightarrow P^\circ \cap \delta L = \neg \emptyset.$$

Tabela A.5 - Relacionamentos ponto-linha

<i>cas o</i>	$P^\circ \cap \delta L$	$P^\circ \cap L^\circ$	<i>relacionamento</i>
1	\emptyset	\emptyset	<i>P disjoint L</i>
2	\emptyset	$\neg \emptyset$	<i>P in L</i>
3	$\neg \emptyset$	\emptyset	<i>P touch L</i>
4	$\neg \emptyset$	$\neg \emptyset$	-

9.4 RELACIONAMENTOS PONTO-PONTO

No caso de dois pontos, deve-se considerar apenas a intersecção de seus interiores, como mostra a tabela 4.A. Adotamos o termo “*overlap*” em lugar de “*in*” por analogia ao caso de relacionamentos linha-linha.

1. P_1 *disjoint* P_2 - dois pontos são disjuntos quando não se interceptam.

$$P_1 \text{ disjoint } P_2 \Leftrightarrow P_1 \cap P_2 = \emptyset.$$

2. P_2 *overlap* P_1 - dois pontos se sobrepõem quando seus interiores se tocam.

$$P_2 \text{ overlap } P_1 \Leftrightarrow (P_1^\circ \cap P_2^\circ = \neg \emptyset).$$

<i>cas o</i>	$P_1^\circ \cap P_2^\circ$	<i>relacionamento</i>
1	\emptyset	$P_1 \text{ disjoint } P_2$
2	$\neg \emptyset$	$P_1 \text{ overlap } P_2$