

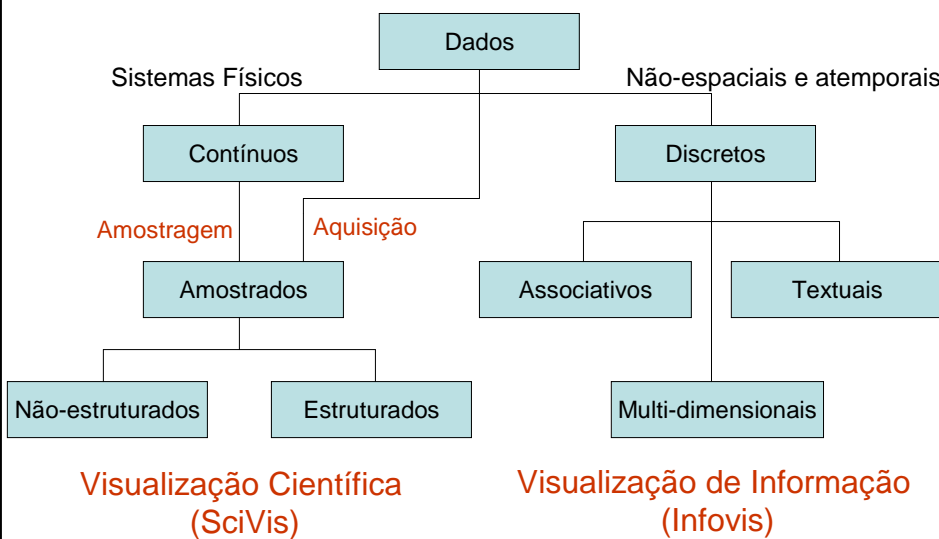
IA369P – Tópicos em Engenharia de Computação VI

Visualização de Informação: Algoritmos

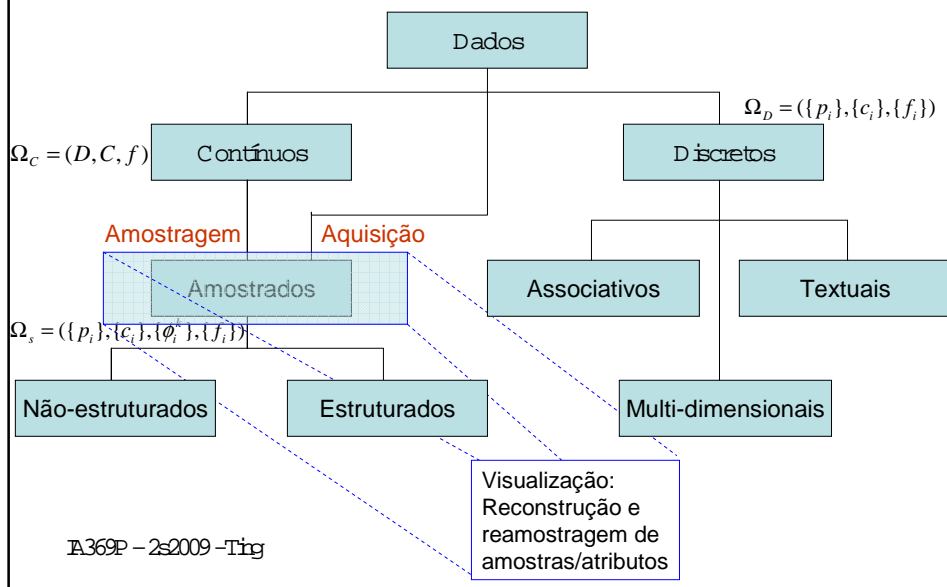
Visualização de Informação

Capítulo 11 do livro-texto Telea
Capítulos 9, 10 do livro-texto Ware

Classificação



R e p r e s e n t a ç ã o d e D a d o s



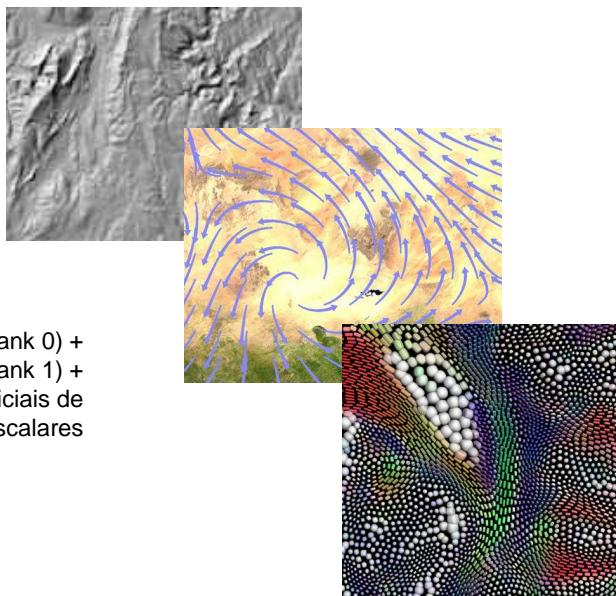
Tipos de Dados Contínuos

Escalares:
 $f(X) \subset R$

Vetores:
 $f(X) \subset R^d$

Tensores:
Escalares (tensores de rank 0) +
vetores (tensores de rank 1) +
outros arranjos matriciais de
escalares

Cores



Exercícios

1. Proponha uma forma para visualizar
 - mobilidade de cardumes de bacalhau no Mar de Norte e nas águas do Ártico
 - organização dos códigos de um aplicativo
 - organização de um acervo de filmes de acordo com a preferência de cada cliente
 - tráfego de dados em uma rede de comunicação

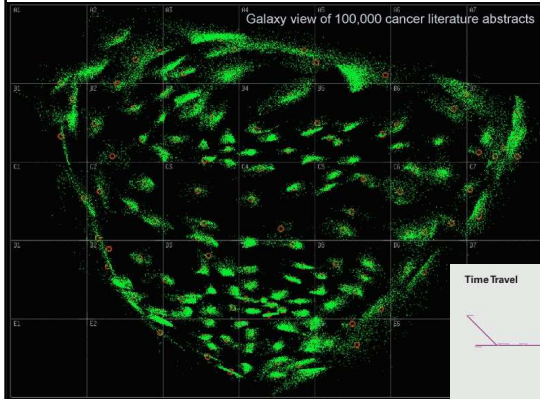
IA369P – 2s2009 - Ting

Informação

- Mobilidade de cardumes de bacalhau no Mar de Norte e nas águas do Ártico
 - Escala espacial
- Organização dos códigos de um aplicativo
 - Escala estrutural
- Organização de um acervo de filmes de acordo com a preferência de cada cliente
 - Escala estrutural
- Tráfego de dados em uma rede de comunicação
 - Escala temporal

IA369P – 2s2009 - Ting

Informação

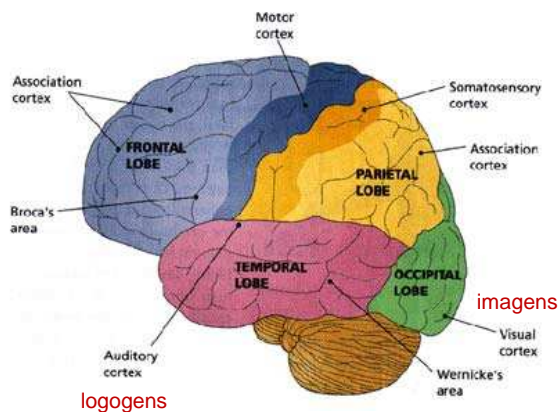


- Dados abstratos.
- Dados desprovidos de forma e de cor.
- Dados não-espaciais

IA369P – 2s2009 - Ting

Informação Visual x Verbal

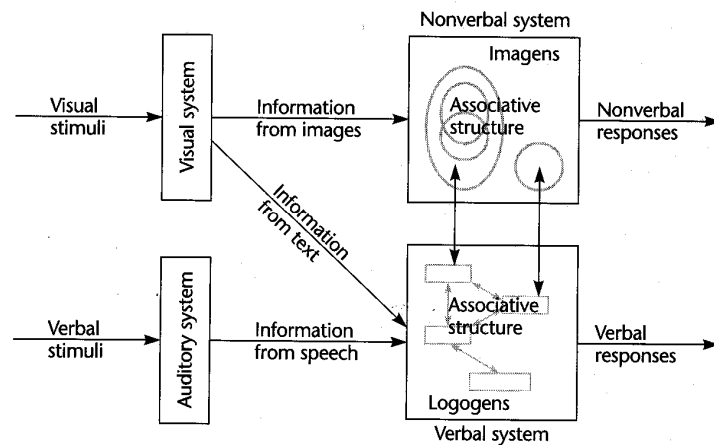
- Dois sistemas sensoriais no cérebro:
 - Córtex visual: processa informações visuais
 - Córtex auditivo (lobo temporal): processa informações auditivas



IA369P – 2s2009 - Ting

Informação Visual x Verbal

- Teoria de Codificação Dual (Paivio, 1987)



IA369P – 2s2009 - Ting

Informação Visual x Verbal

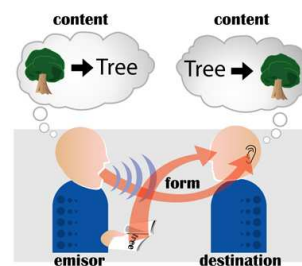
Visual Estática



- é compreendida de forma paralela;
- pode ser multi-cultural;
- é “nata”;
- é mais adequada para mostrar a topologia de um sistema (relações entre elementos), localizações espaciais e detalhes na superfície visível.

IA369P – 2s2009 - Ting

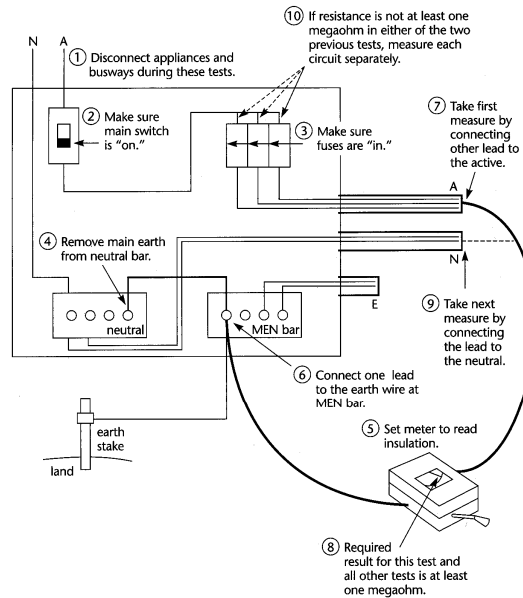
Verbal



- é dinâmica e seqüencial no tempo;
- é a mais utilizada;
- é ensinada desde os primeiros dias de vida;
- é mais adequada para conceitos abstratos, detalhes de uma informação abstrata, informações conjuntivas, concessivas, causais, temporais e lógicas.

Informação Visual x Verbal

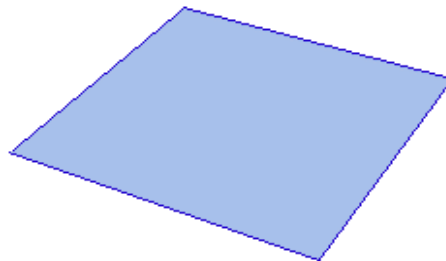
Instruções para teste de um sistema elétrico



IA369P – 2s2009 - Ting

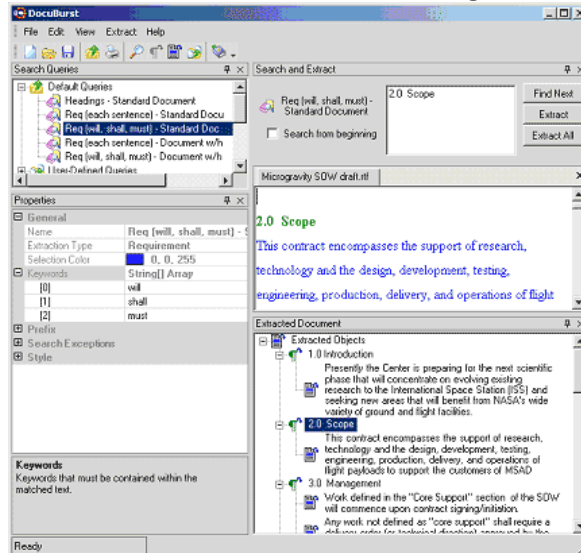
Informação Visual x Verbal

- **Animações** são úteis para mostrar um conceito dinâmico, envolvendo movimentos complexos, causa-efeito, passos de uma transformação, instruções.



IA369P – 2s2009 - Ting

Informação



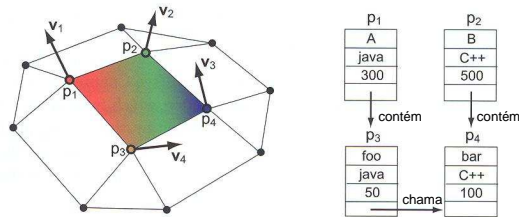
DocuBurst

(Fonte:

http://www.tbe.com/docu_burst.html

IA369P – 2s2009 - Ting

Tipos de Dados

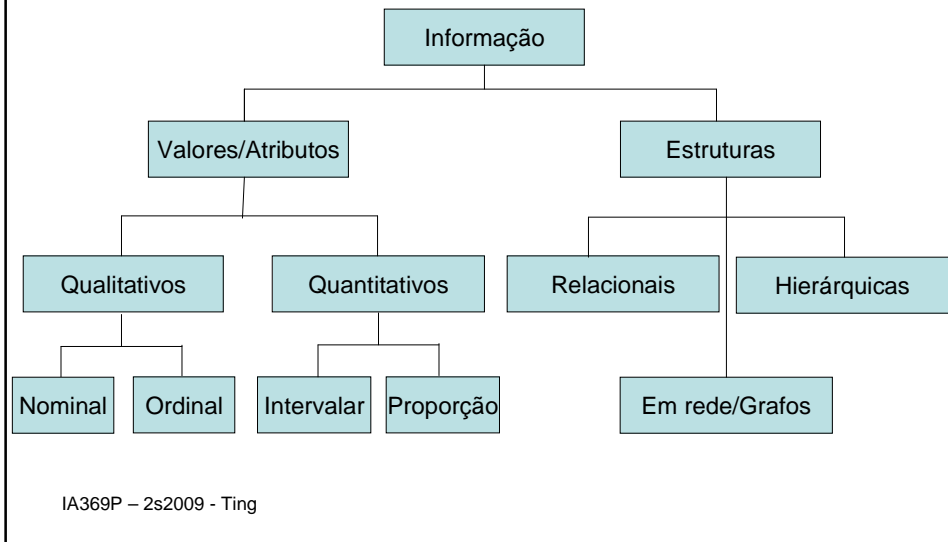


	SciVis	InfoVis
Domínio	Espacial, temporal	não-espacial, atemporal, estrutural
Tipos de valores/atributos	numérico	nominal, ordinal, numérico
Dados	amostras	n-tuplas de atributos
“Células”	interpolação	relação
Interpolação	contínua por parte	Pode não existir

IA369P – 2s2009 - Ting

Informação

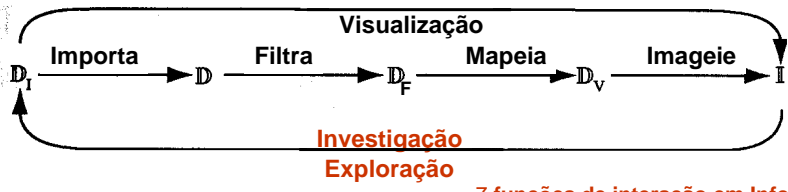
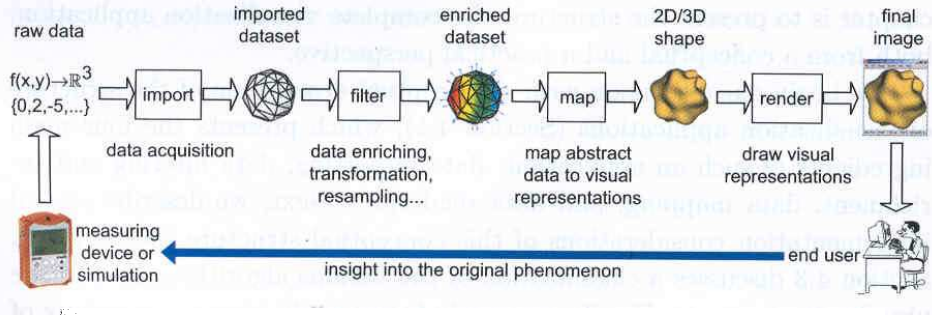
Uma Classificação



Tipos de Atributos

Atributos	Domínio	Operações	Exemplos
Nominal	Conjunto não-ordenado	Comparações (igualdade)	nomes de objetos, números de identificação
Ordinal	Conjunto ordenado	Comparações (igualdade, maior, menor)	classificação de uma avaliação
discretos	Domínio de números inteiros	Operações sobre inteiros	Linhas de um programa, quantidade de caracteres em um texto
Intervalar + proporção	Domínio de números reais	Operações sobre reais	tempo, altura, distância

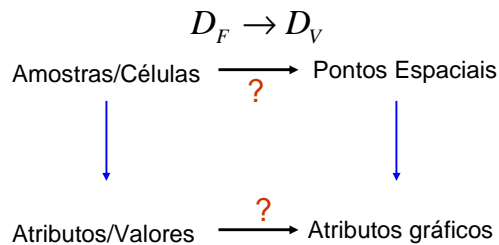
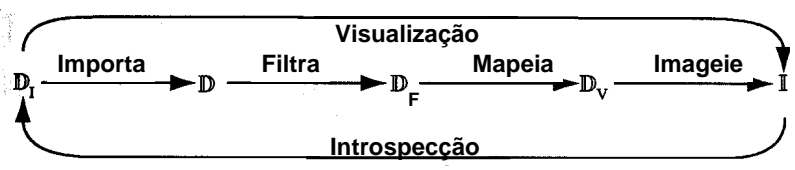
Modelo Conceitual



IA369P – 2s2009 - Ting

7 funções de interação em InfoVis:
 overview, zoom, filter, details-on-demand, relate, history, extract

Modelo Conceitual



Uma diretriz:

- Adotar convenções consagradas na área
- Maximizar o reuso das ferramentas existentes
- Maximizar a ocupação dos *pixels*

IA369P – 2s2009 - Ting

Exercícios

1. O que você entende por informação verbal e informação visual? Qual delas estamos mais aptos para processar?
2. Compare InfoVis e SciVis em termos de domínios, tipos de atributos e formas de interpolação.
3. Por que, em termos de visualização, InfoVis é um tópico mais complexo do que SciVis?
4. Por que em InfoVis é comum combinar texto e imagens para representar uma informação?

IA369P – 2s2009 - Ting

Visualização de Dados Tabulares

Dados tabulares = dados estruturáveis em linhas e colunas.

rows of equal date value

first sort on name

then sort on date

and then on time

evolution icons

id	category	name	date	time	open	high	low	close
636	sf	SIF1	2004-11-29	13:00	0.800000	0.800000	0.800000	0.800000
635	sf	SIF1	2004-11-29	14:00	0.800000	0.800000	0.800000	0.800000
633	sf	SIF1	2004-11-29	15:00	0.795000	0.795000	0.795000	0.795000
630	sf	SIF1	2004-11-30	14:00	0.795000	0.795000	0.795000	0.795000
632	sf	SIF1	2004-11-30	12:00	0.800000	0.800000	0.795000	0.795000
621	sf	SIF1	2004-11-30	13:00	0.795000	0.795000	0.795000	0.795000
628	sf	SIF1	2004-11-30	15:00	0.795000	0.795000	0.795000	0.795000
627	sf	SIF1	2005-00-02	12:00	0.795000	0.795000	0.790000	0.790000
626	sf	SIF1	2005-00-02	13:00	0.790000	0.795000	0.790000	0.795000
625	sf	SIF1	2005-00-02	14:00	0.795000	0.795000	0.795000	0.795000
624	sf	SIF1	2005-00-02	15:00	0.800000	0.800000	0.800000	0.800000
620	sf	SIF1	2005-00-03	15:00	0.795000	0.795000	0.795000	0.795000
622	sf	SIF1	2005-00-03	12:00	0.795000	0.795000	0.795000	0.795000
621	sf	SIF1	2005-00-03	13:00	0.795000	0.795000	0.795000	0.795000
619	sf	SIF1	2005-00-03	14:00	0.795000	0.795000	0.795000	0.795000
618	sf	SIF1	2005-00-06	15:00	0.795000	0.795000	0.795000	0.795000
617	sf	SIF1	2005-00-06	11:00	0.790000	0.790000	0.790000	0.790000
616	sf	SIF1	2005-00-06	15:00	0.795000	0.795000	0.795000	0.795000
615	sf	SIF1	2005-00-06	12:00	0.795000	0.795000	0.795000	0.795000
614	sf	SIF1	2005-00-06	13:00	0.795000	0.795000	0.795000	0.795000
613	sf	SIF1	2005-00-06	14:00	0.795000	0.795000	0.795000	0.795000
609	sf	SIF1	2005-00-07	14:00	0.790000	0.790000	0.790000	0.790000
612	sf	SIF1	2005-00-07	11:00	0.795000	0.795000	0.795000	0.795000
611	sf	SIF1	2005-00-07	12:00	0.795000	0.795000	0.795000	0.795000
610	sf	SIF1	2005-00-07	13:00	0.790000	0.790000	0.790000	0.790000
608	sf	SIF1	2005-00-07	15:00	0.790000	0.790000	0.790000	0.790000
606	sf	SIF1	2005-00-08	12:00	0.790000	0.795000	0.795000	0.795000
607	sf	SIF1	2005-00-08	12:00	0.790000	0.790000	0.790000	0.790000
605	sf	SIF1	2005-00-08	14:00	0.795000	0.795000	0.795000	0.795000

Usado de cores para reforçar dados de um mesmo dia

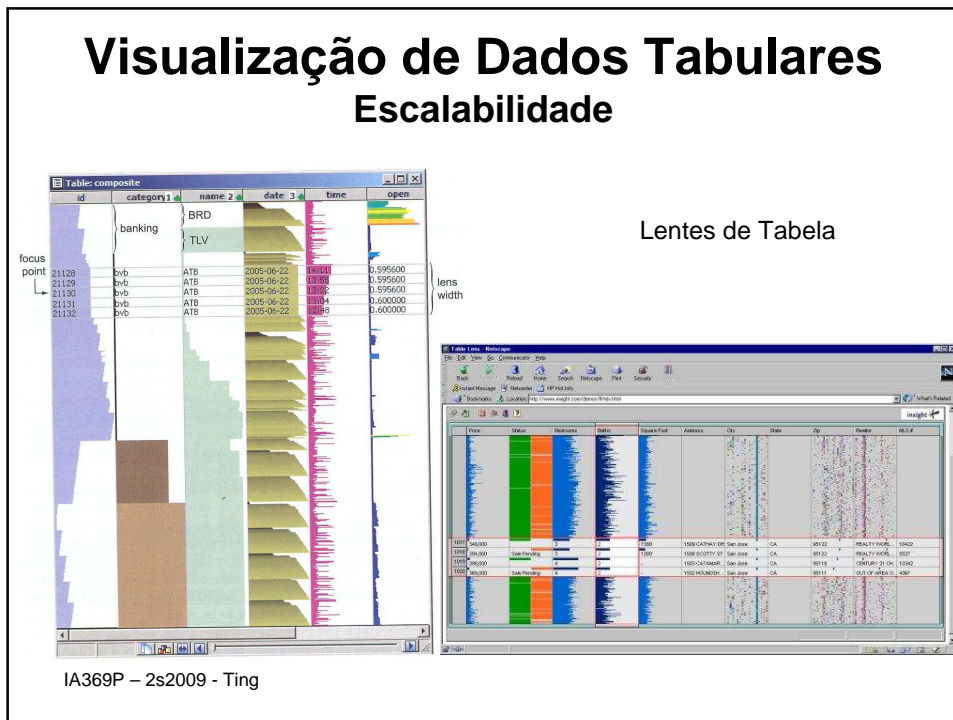
Gráficos de barras

attribute bar graphs

zoom slider

IA369P – 2s2009 - Ting

Visualização de Dados Tabulares Escalabilidade

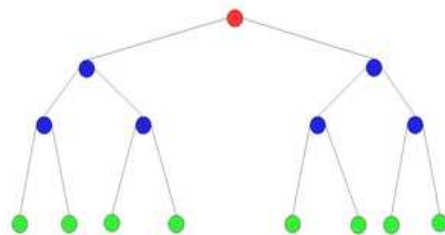


Visualização de Dados Relacionais

- Um conjunto de pontos (vértices/nós) ligados por retas (arestas/arcos).
- Problemas reais:
 - Cidades e estradas em um mapa
 - Tarefas e custos de execução em um planejamento de projetos
 - Terminais/*switches* e cabos de conexão em uma rede
 - Pessoas e frequência de contatos em uma rede social
- Estruturas de dados
 - Estrutura de Árvore
 - Grafo

Árvores

- Uma organização de dados que tem as características topológicas de uma árvore, em que os dados são relacionados de forma hierárquica a partir de um elemento principal (raiz) até as folhas.
- Informação de natureza hierárquica:
 - Sistema de domínio de nomes em redes de computadores
 - Sistema de pastas em computadores
 - Organograma
 - Árvore genealógica
- Algoritmos:
 - Árvore
 - *treemaps*



IA369P – 2s2009 - Ting

Visualização de Árvores

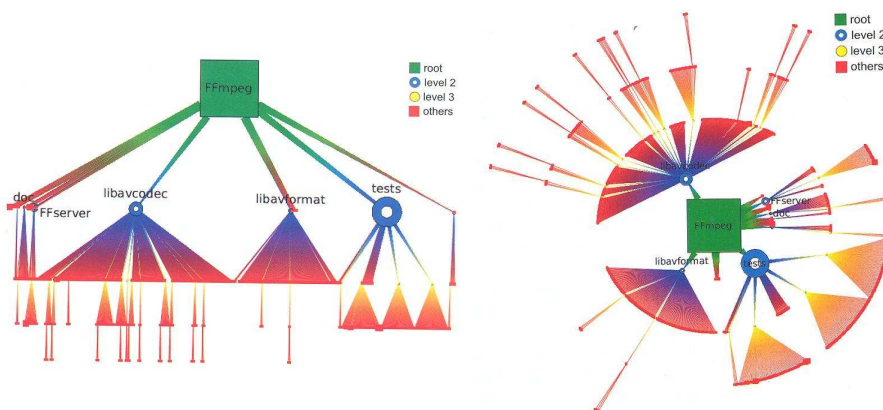


Diagrama em árvore

- Nível → coordenada y
- Largura → coordenada x

Diagrama radial

- Nível → raio
- Largura → ângulo

IA369P – 2s2009 - Ting

Visualização de Árvores

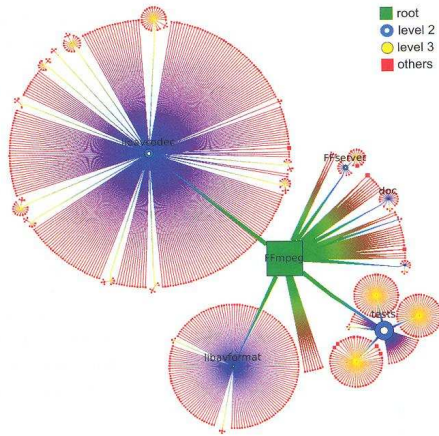


Diagrama de bolha

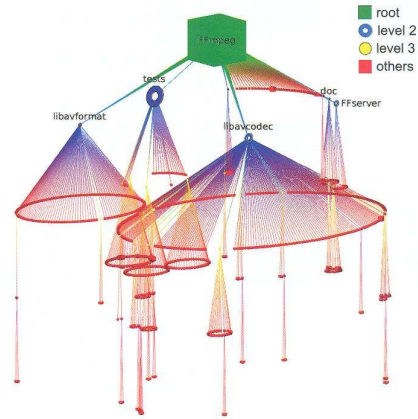
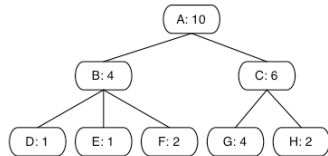


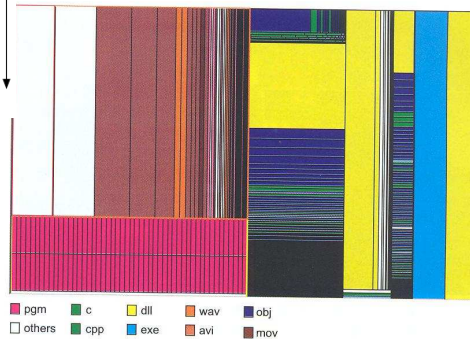
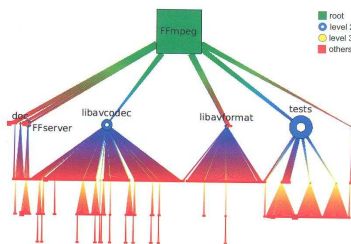
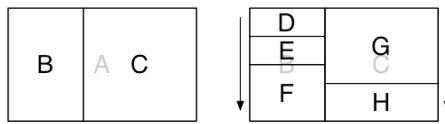
Diagrama em cone

IA369P – 2s2009 - Ting

Visualização de Árvores



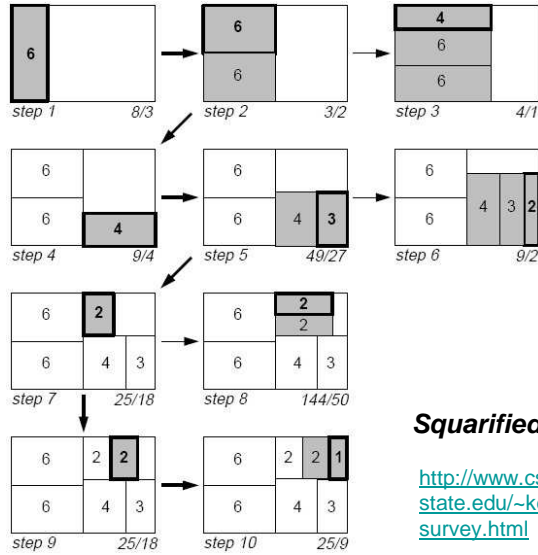
Schneiderman (1992): *Slice-and-dice*



IA369P – 2s2009 - Ting

<http://www.cs.umd.edu/hcil/treemap-history/index.shtml>

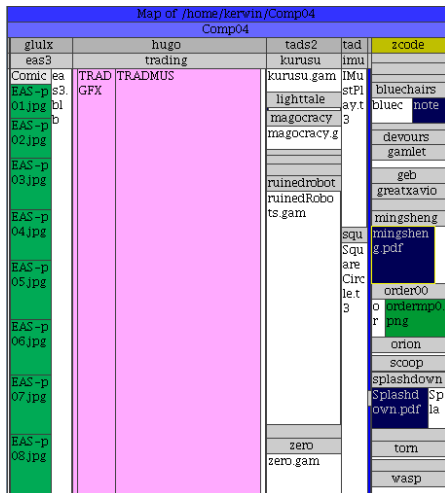
Visualização de Árvores



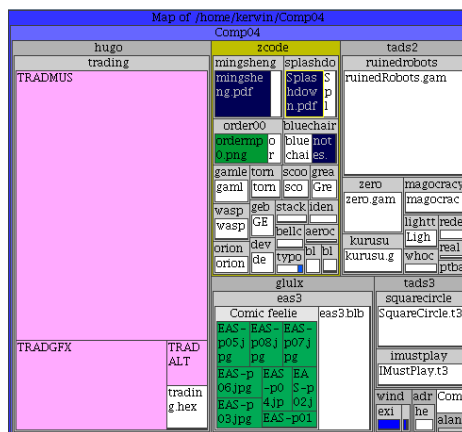
Squarified treemap layout

<http://www.cse.ohio-state.edu/~kerwin/treemap-survey.html>

Visualização de Árvores

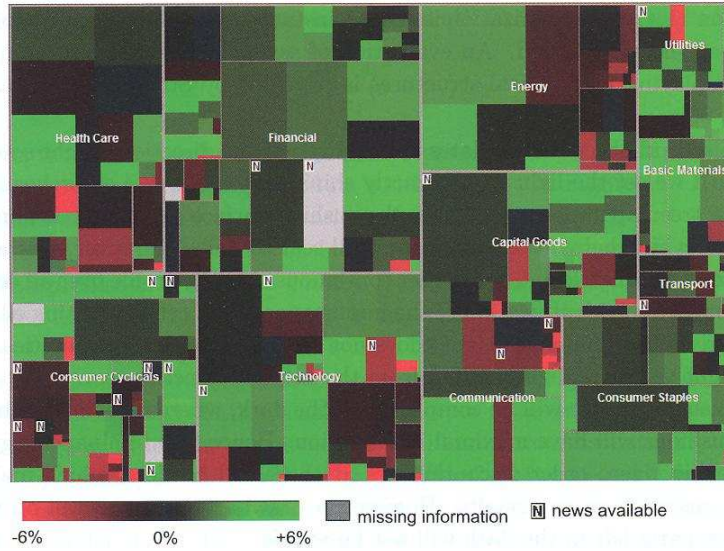


Slice-and-dice



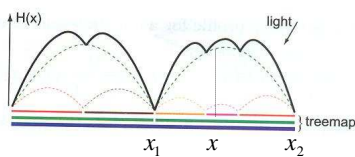
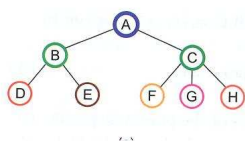
Squarified treemap

Visualização de Árvores



IA369P – 2s2009 - Ting

Visualização de Árvores



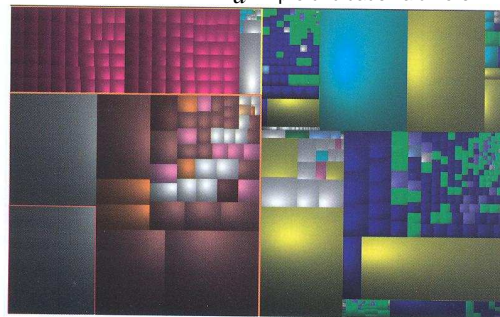
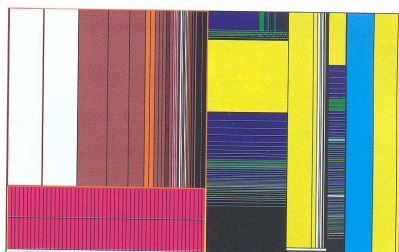
$$H(x) = \frac{4k}{x_2 - x_1} (x - x_1)(x_2 - x)$$

$$k = f^d K$$

$$f, K \in [0, 1]$$

$$d \rightarrow \text{profundidade na árvore}$$

Cushion treemap



IA369P – 2s2009 - Ting

http://w3.win.tue.nl/nl/onderzoek/onderzoek_informatica/visualization/sequoiaview/

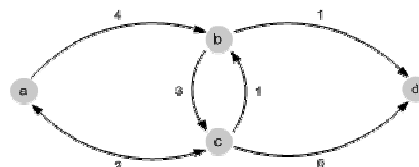
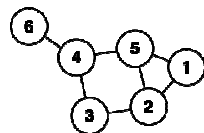
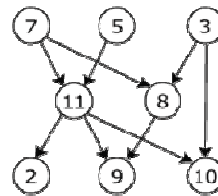
Visualização de Árvores

- Pontos-e-Varetas/Vértices-e-arestas
 - Minimizar sobreposições
 - Minimizar comprimento das arestas
 - Manter razão-de-aspecto em torno de 1
 - Escalabilidade em legibilidade e em recursos computacionais
 - Relações são facilmente discerníveis
 - Maximizar clusterização
- *Treemaps*: uma técnica de visualização de dados hierárquicos condicionada a um espaço pré-estabelecido
 - Assegurar a separabilidade entre os nós
 - Assegurar a visibilidade de cada nó
 - Otimizar a ocupação dos *pixels*
 - Manter razão-de-aspecto em torno de 1

IA369P – 2s2009 - Ting

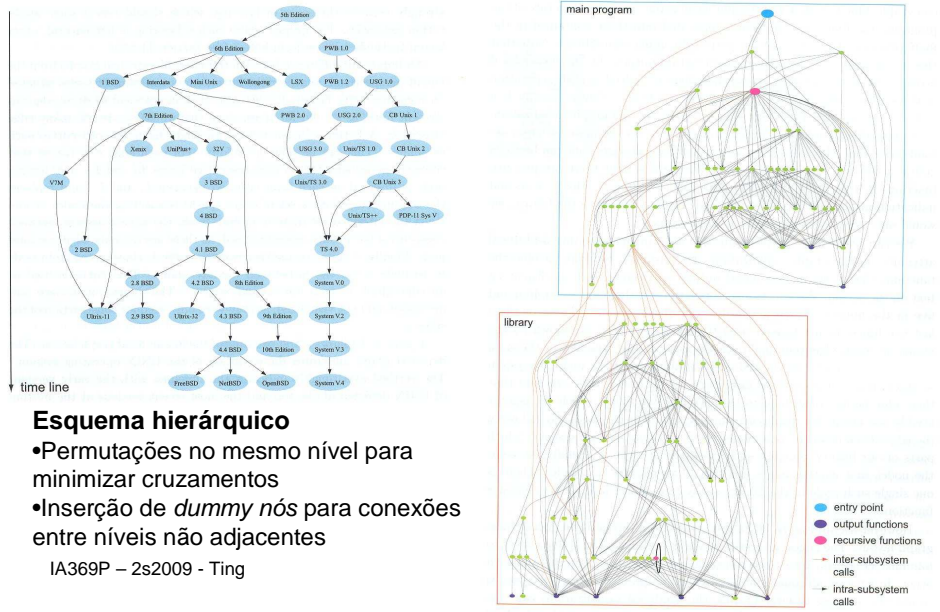
Grafos

- Generalização de estrutura de dados relacionais.
- $G=\{N,E\}$, onde $N=\{n_i\}$ é um conjunto de nós e $E=\{e_i\}$, um conjunto de arestas que conectam os nós $e_i=(n_j,n_k)$.
- Tipos de grafos:
 - Com laços ou sem laços
 - Dirigidos e não-dirigidos
- Algoritmos
 - Hierárquicos
 - Baseados em forças



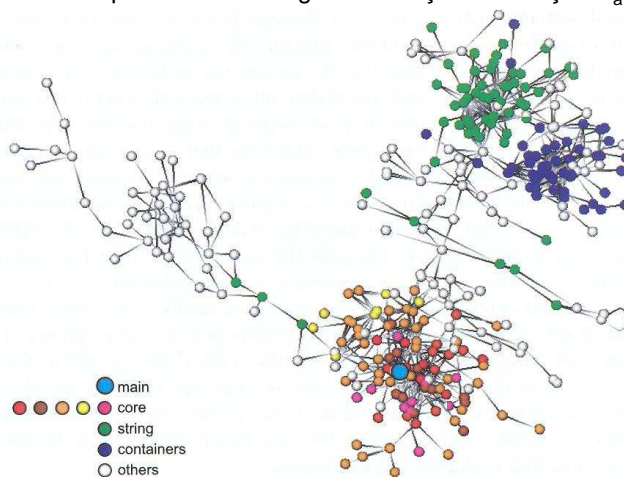
IA369P – 2s2009 - Ting

Visualização de Grafos



Visualização de Grafos

Parâmetros de qualidade de um grafo → forças de atração F_a e repulsão F_r



Visualização de Grafos

- Fruchterman e Reingold, 1991

$$F_a(n_i, n_j) = \frac{|p_i - p_j|^2}{k}$$

$$F_r(n_i, n_j) = -\frac{k^2}{|p_i - p_j|}$$

$$k = \frac{\sqrt{A}}{|N|}$$

A: área de desenho
N: número de nós

- Eades, 1984

$$F_a(n_i, n_j) = k \log |p_i - p_j|$$

$$F_r(n_i, n_j) = -\frac{k^2}{|p_i - p_j|^2}$$

- Kamada e Kawai, 1989

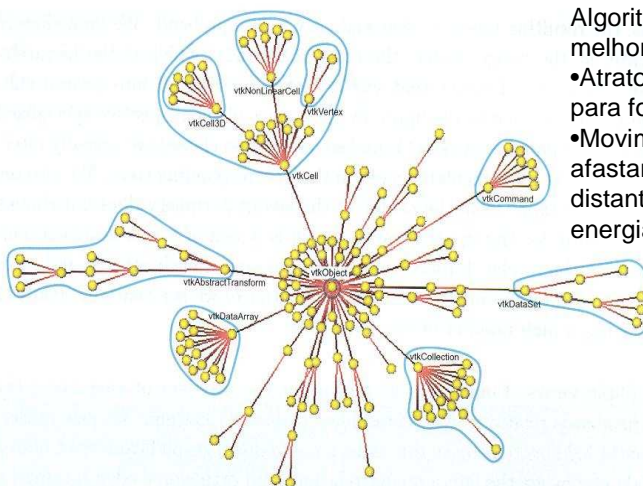
$$\mathcal{E} = \frac{1}{2} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} c_{ij} (|p_i - p_j| - d_{ij})^2$$

$$c_{ij} = \begin{cases} 1 & \text{Existe aresta entre } n_i \text{ e } n_j \\ 0 & \end{cases}$$

d_{ij} Menor distância entre n_i e n_j

IA369P – 2s2009 - Ting

Visualização de Grafos

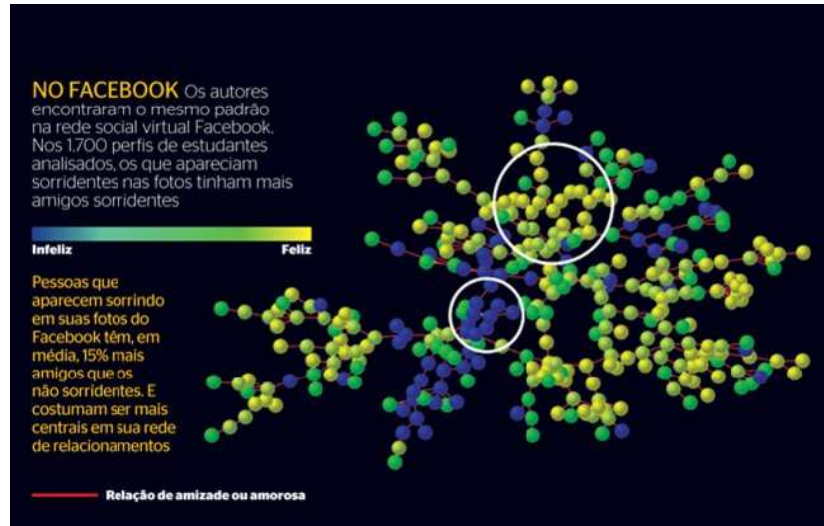


Algoritmo baseado em forças melhorado:

- Atrator fictício no baricentro para forçar agrupamento
- Movimentos aleatórios para afastar do mínimo local, distante do mínimo global de energia.

IA369P – 2s2009 - Ting

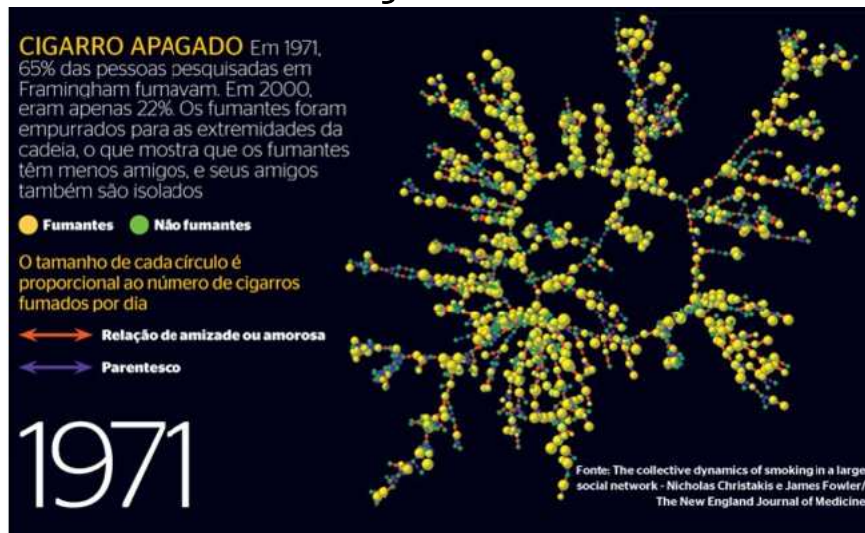
Visualização de Grafos



IA369P – 2s2009 - Ting

<http://revistaepoca.globo.com/Revista/Epoca/0,,EM1100540-15224-1.00-O+PODER+DA+AMIZADE.html>

Visualização de Grafos



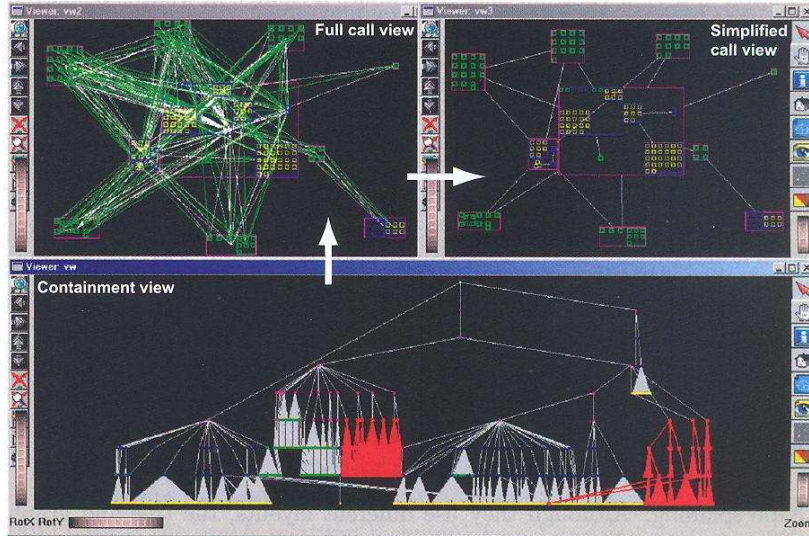
IA369P – 2s2009 - Ting

<http://revistaepoca.globo.com/Revista/Epoca/0,,EM1100540-15224-1.00-O+PODER+DA+AMIZADE.html>

Visualização em multi-vistas

Details

Details

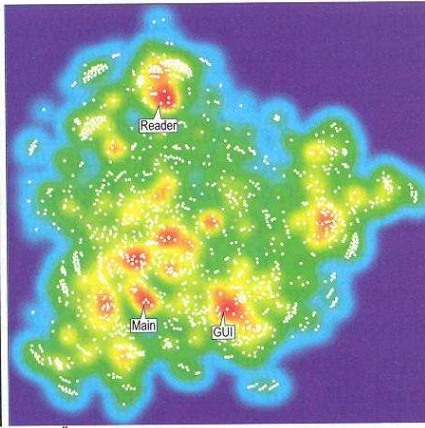
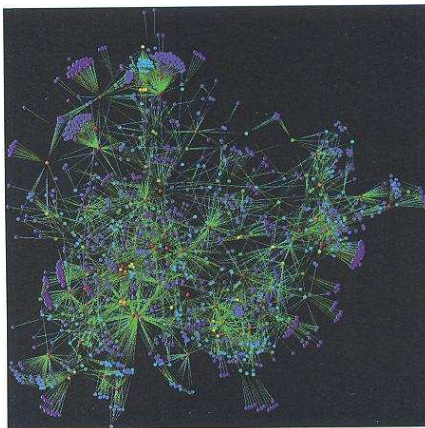


IA369P – 2s2009 - Ting

Overview

Visualização de Grafos

Simplificação de Grafos → Técnica de *Splatting*



$$f(p) = \sum_{i=1}^n f_i \Phi_i(p)$$

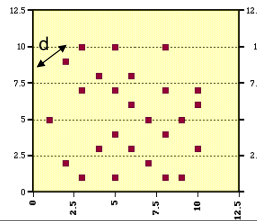
f_i fator de contribuição do nó n_i
 $\Phi_i(p) = e^{-k|p-p_i|^2}$ Função de base radial

IA369P – 2s2009 - Ting

Visualização de Dados Multidimensionais

- 2 atributos → (x,y)
- 3 atributos → (x,y,z)
- 4 atributos → (x,y,z,cor)
- 5 atributos → (x,y,z,matiz,luminância)
- k atributos → vetores de dimensão k → distância entre vetores

$$d(a_1, \dots, a_k, b_1, \dots, b_k) = \sqrt{\sum_{i=1}^k \delta(a_i, b_i)^2}$$

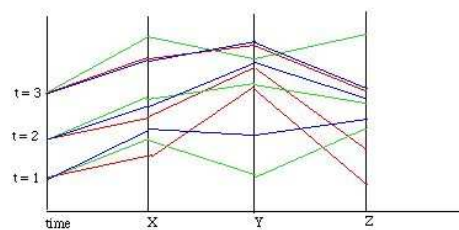
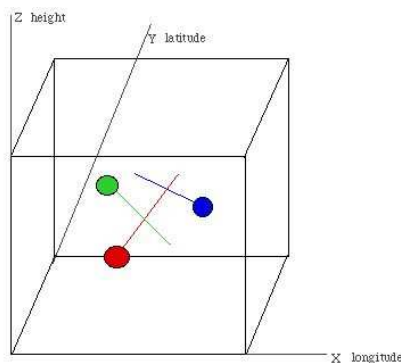


IA369P – 2s2009 - Ting

Há perda de informação!!

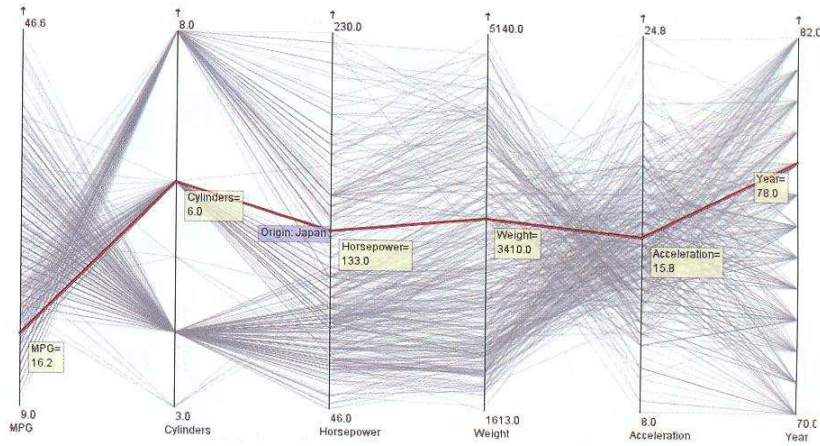
Visualização de Dados Multidimensionais

Localização espacial de 3 aeronaves



IA369P – 2s2009 - Ting

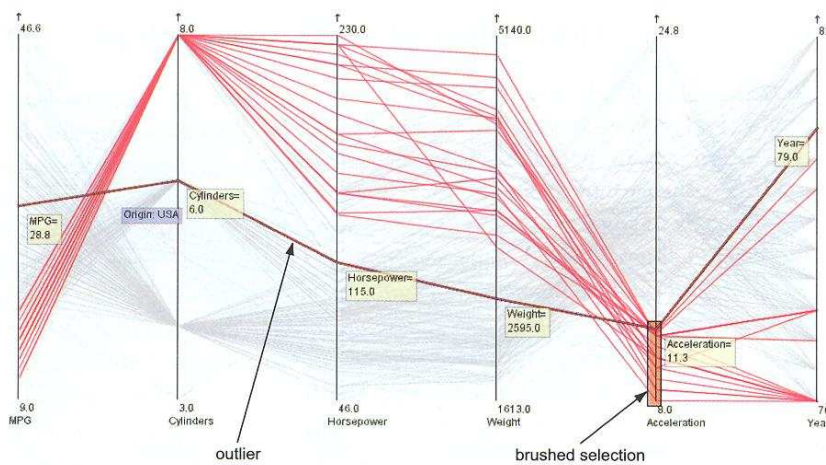
Visualização de Dados Multidimensionais



IA369P – 2s2009 - Ting

Coordenadas Paralelas

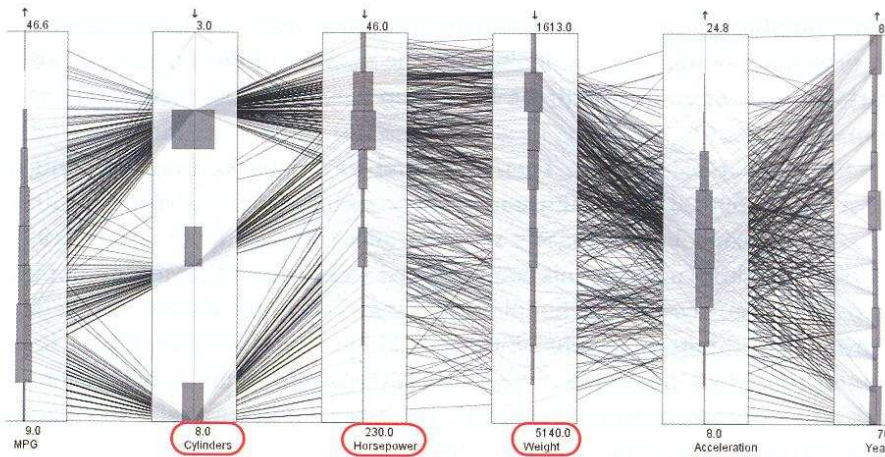
Visualização de Dados Multidimensionais



IA369P – 2s2009 - Ting

Interação: *brushing*

Visualização de Dados Multidimensionais



IA369P – 2s2009 - Ting

Mais artifícios para aprimorar legibilidade e interatividade:

- Inversão na ordenação
- Sobreposição de histogramas para cada atributo

Exercícios

1. Supondo conhecidas as coordenadas (x,y) de cada nó e as relações entre eles Escreva um pseudo-código de desenho de uma árvore de forma que o número de cruzamentos entre as arestas menor possível. Defenda a heurística que você adotou.
2. Explique sucintamente o princípio básico de um algoritmo de *treemap*. Compare as vantagens e as desvantagens desta forma de visualização em relação a de árvore.
3. Em que difere essencialmente as árvores dos grafos para visualizar uma dada informação? Até qual extensão um algoritmo de árvore pode ser adaptado para visualizar os grafos?
4. Explique sucintamente um algoritmo de desenho de grafos baseado em forças de atração e de repulsão. Quais são as vantagens e desvantagens deste algoritmo?

IA369P – 2s2009 - Ting

Texto

- Texto = Conteúdo + Estrutura + Meta-dados
 - Conteúdo: diferentes tipos de informação (abstratos e concretos)
 - Estrutura: organização do texto (parágrafos, seções, capítulos, etc.)
 - Meta-dados: descrição do documento (autor, edição, palavras-chave, etc.)
- Analisador de textos:
 - Análise léxica
 - Análise sintática
 - Análise semântica
 - Processamento de linguagem natural
- Natureza de textos:
 - narrativo
 - lírico
 - dramático
 - *Software*: códigos de programa

IA369P – 2s2009 - Ting

Informação Visual x Verbal

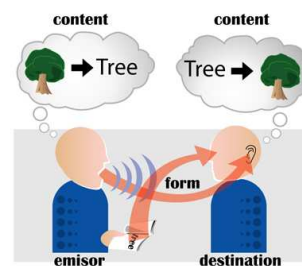
Visual Estática



- é compreendida de forma paralela;
- pode ser multi-cultural;
- é “nata”;
- é mais adequada para mostrar a topologia de um sistema (relações entre elementos), localizações espaciais e detalhes na superfície visível.

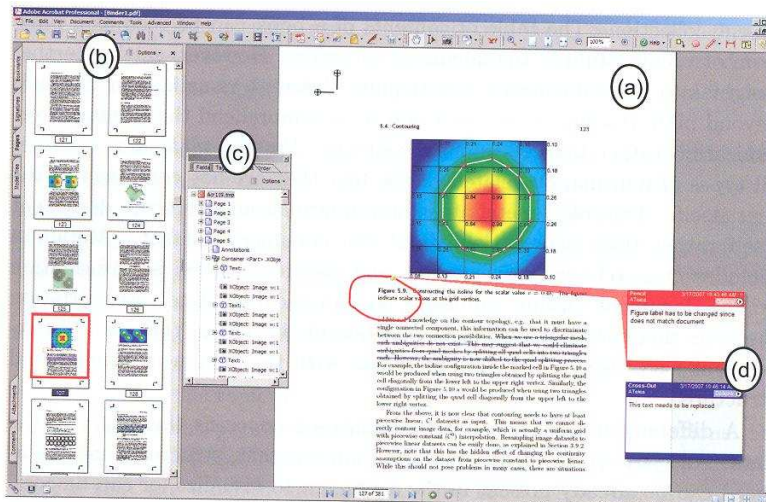
IA369P – 2s2009 - Ting

Verbal



- é dinâmica e seqüencial no tempo;
- é a mais utilizada;
- é ensinada desde os primeiros dias de vida;
- é mais adequada para conceitos abstratos, detalhes de uma informação abstrata, informações conjuntivas, concessivas, causais, temporais e lógicas.

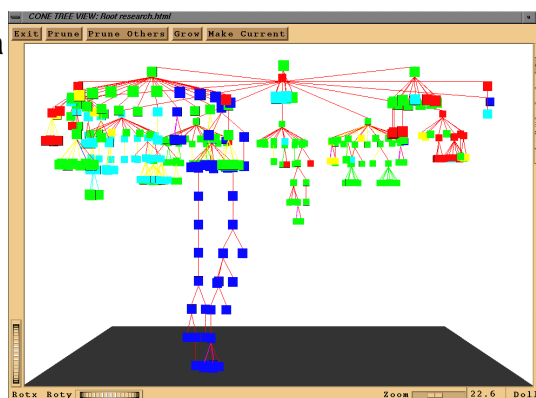
Visualização de Conteúdo



IA369P – 2s2009 - Ting

Visualização de Software

- Flexibilidade para adequar à sintaxe da linguagem de programação selecionada
- Escalabilidade
- Informação relacional
- Informação hierárquica
- Volume de “meta-dados”



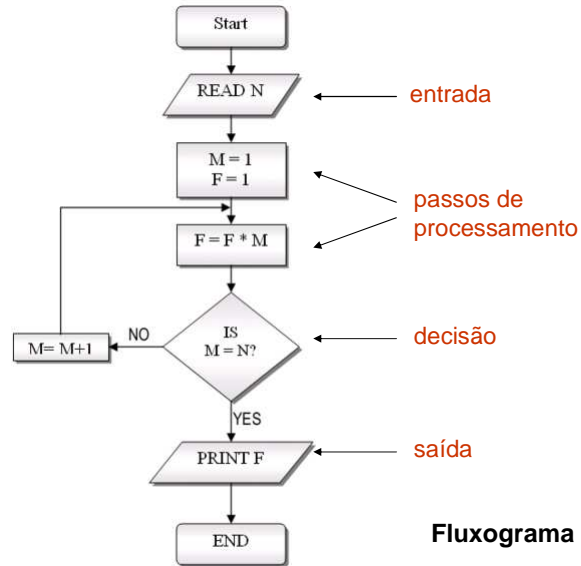
IA369P – 2s2009 - Ting

Visualização de Software

Fatorial de N

```
BEGIN  
Input N  
M=F=1  
WHILE (M ≤ N) DO  
  F = F*M  
  M = M+1  
ENDWHILE  
Display F  
END
```

Pseudo-código

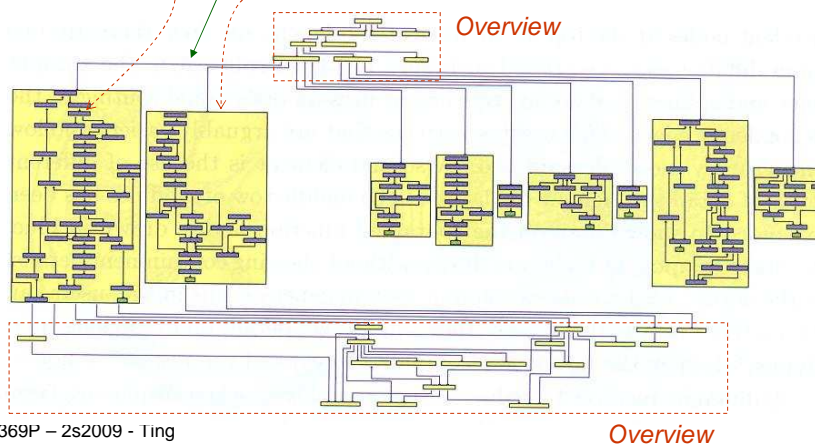


Fluxograma

IA369P – 2s2009 - Ting

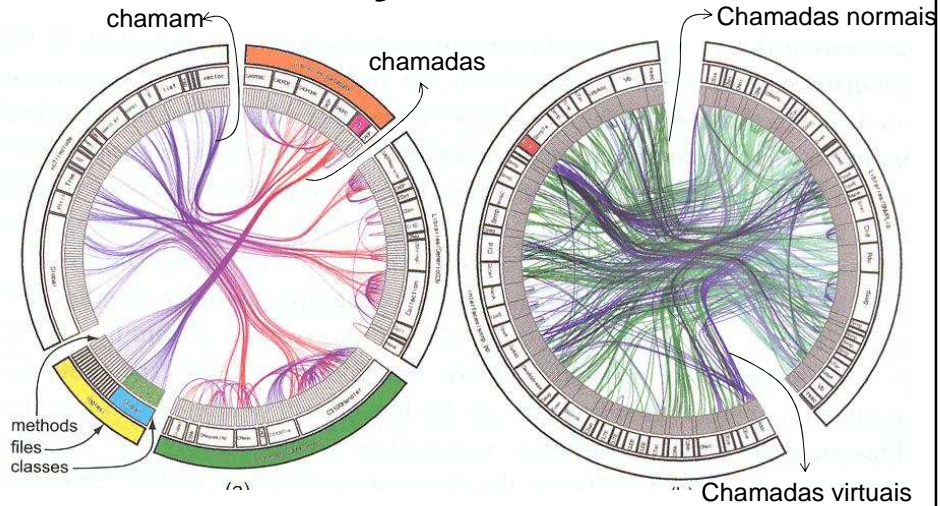
Visualização de Software

- Representação de códigos orientados a objetos
 - Nós: métodos, classes e arquivos
 - Relações: dependências e pertinências



IA369P – 2s2009 - Ting

Visualização de Software



Código bem estruturado

Código mal estruturado

IA369P – 2s2009 - Ting

Visualização de Software

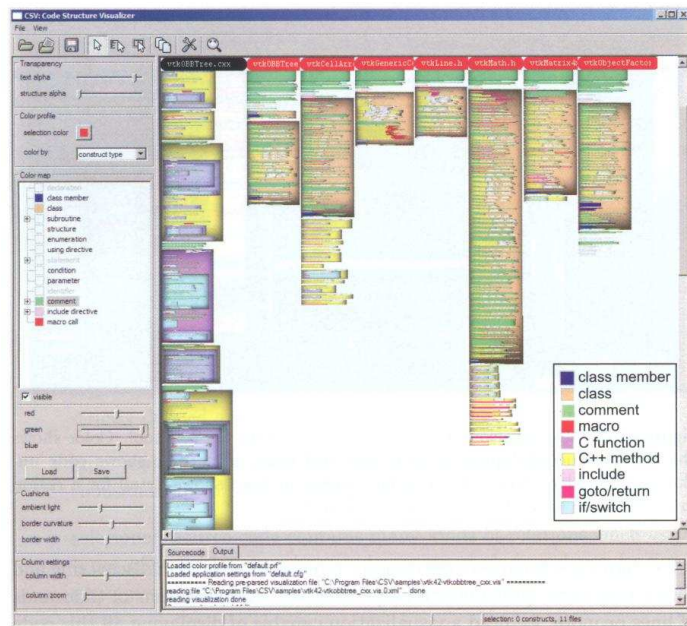


SeeSoft
 Projeto de Software
 (Fonte: http://www.cc.gatech.edu/classes/cs7390_98_winter/reports/realsys/seesoft.html)

IA369P – 2s2009 - Ting

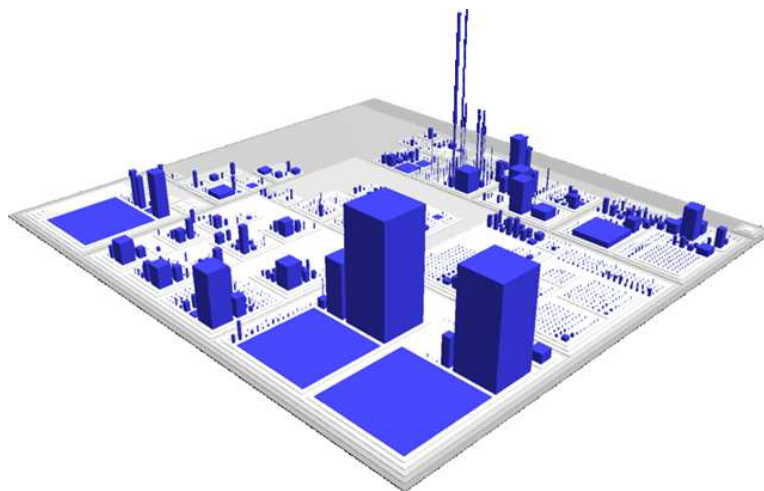
Visualização de Software

E outras relações,
além das
chamadas e
pertinências?



IA369P – 2s2009 - Ting

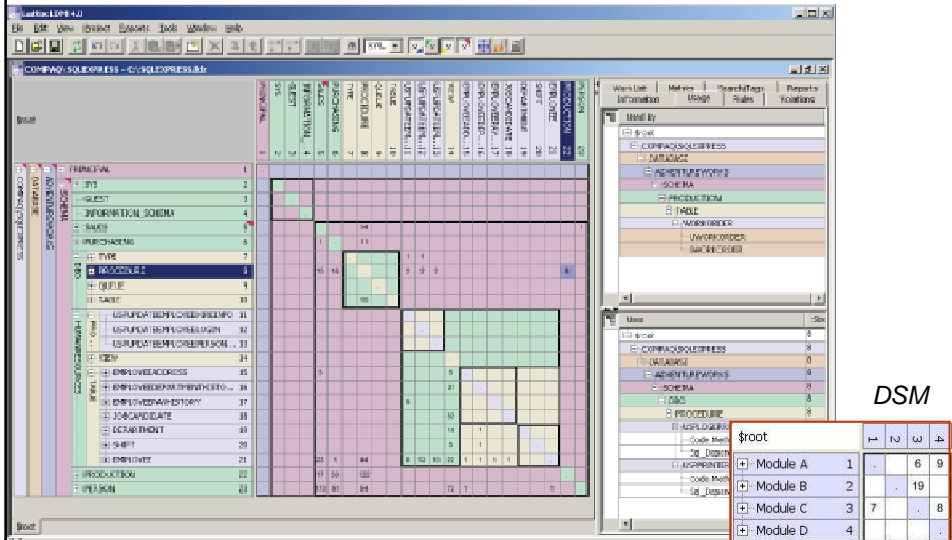
Visualização de Software



IA369P – 2s2009 - Ting

<http://www.inf.unisi.ch/phd/wettel/codacity.html>

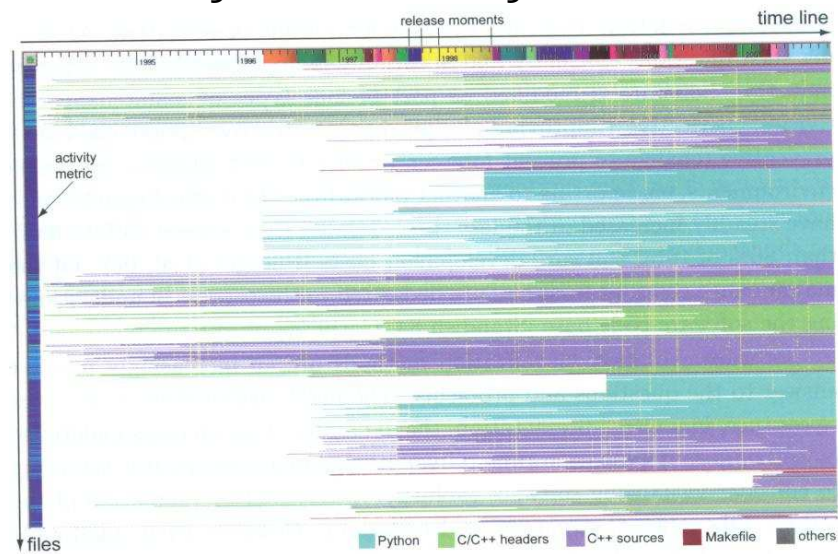
Visualização de Software



IA369P – 2s2009 - Ting

<http://www.lattix.com/node/32>

Visualização de Evolução de Software



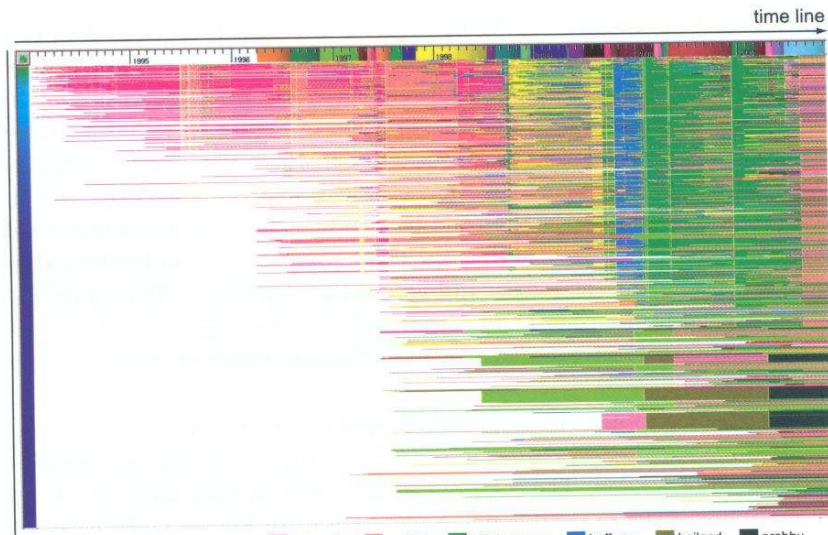
IA369P – 2s2009 - Ting

Visualização de Evolução de Software

- Recuperação de versões anteriores
- Comparação com versões anteriores
- **Manutenção**: aperfeiçoamento, correção, prevenção e adaptação
- Visualização para
 - Visão geral da estruturação do software
 - Modificações feitas ao longo de um período: quem e onde.
 - Diferenças entre as versões
 - Previsão da tendência de alterações no desenvolvimento

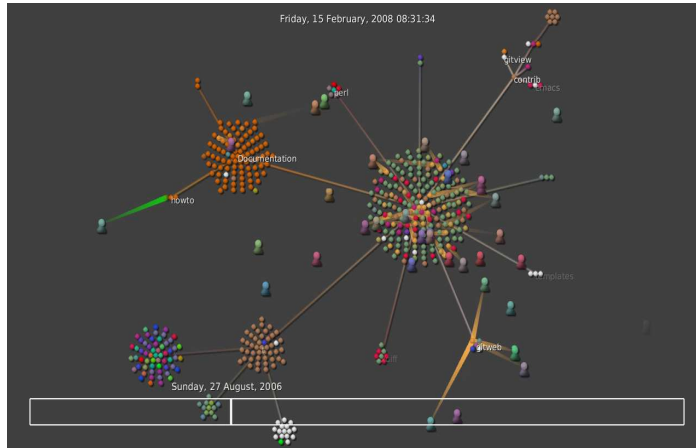
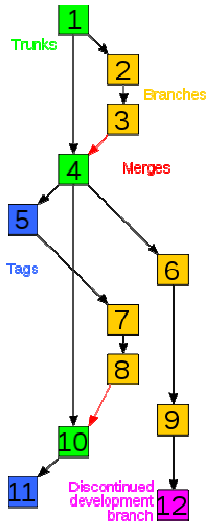
IA369P – 2s2009 - Ting

Visualização de Evolução de Software



IA369P – 2s2009 - Ting

Visualização de Evolução de Software



IA369P – 2s2009 - Ting

<http://github.com/acaudwell/Gource>

Exercícios

1. A visualização de um texto pode ocorrer em diferentes níveis. Qual nível é ainda um desafio para a área de pesquisa de visualização? Qual é a barreira? Como você acha que se pode transpor esta barreira?
2. Em que difere um documento textual de um *software*? Por que o segundo apresenta menos problemas que o primeiro na visualização?
3. O que você entende por manutenção de *software*?
4. Por que a visualização da evolução de um *software* tem sido considerado como uma ferramenta importante para manutenção de *software*?

IA369P – 2s2009 - Ting