

INE5430

Inteligência Artificial

Tópico:

♦ Raciocínio Baseado em Casos
Baseado no material do prof. Luis Otavio Alvares

Raciocínio Baseado em Casos *CBR (Case-Based Reasoning)*

- Paradigma para resolução de problemas
- Ao resolver um novo problema, considera a solução utilizada em problemas similares
- Uma abordagem incremental



CBR (Case-Based Reasoning)

“Um sistema de CBR resolve problemas por adaptar soluções que foram utilizadas para resolver problemas anteriores.”

Riesbeck & Schank, 1989

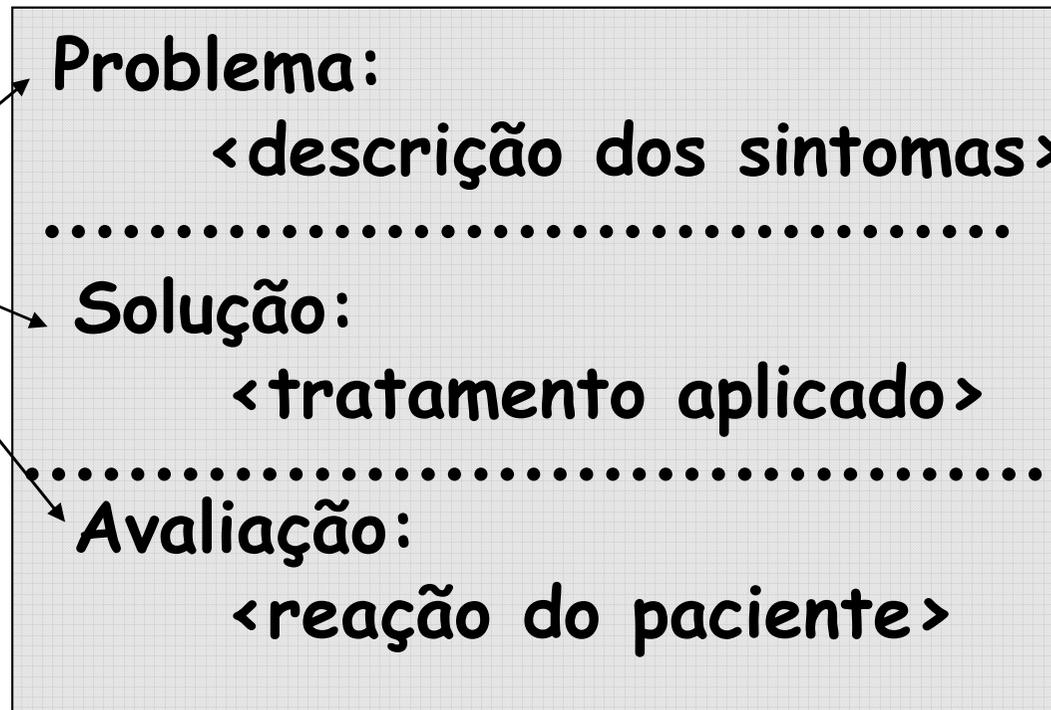
Caso

É uma descrição completa do problema do domínio, com a respectiva solução aplicada, mais uma avaliação da eficácia desta solução

Exemplo de Caso

- ◆ Sistema para diagnóstico de doença
- ◆ Exemplo de caso :

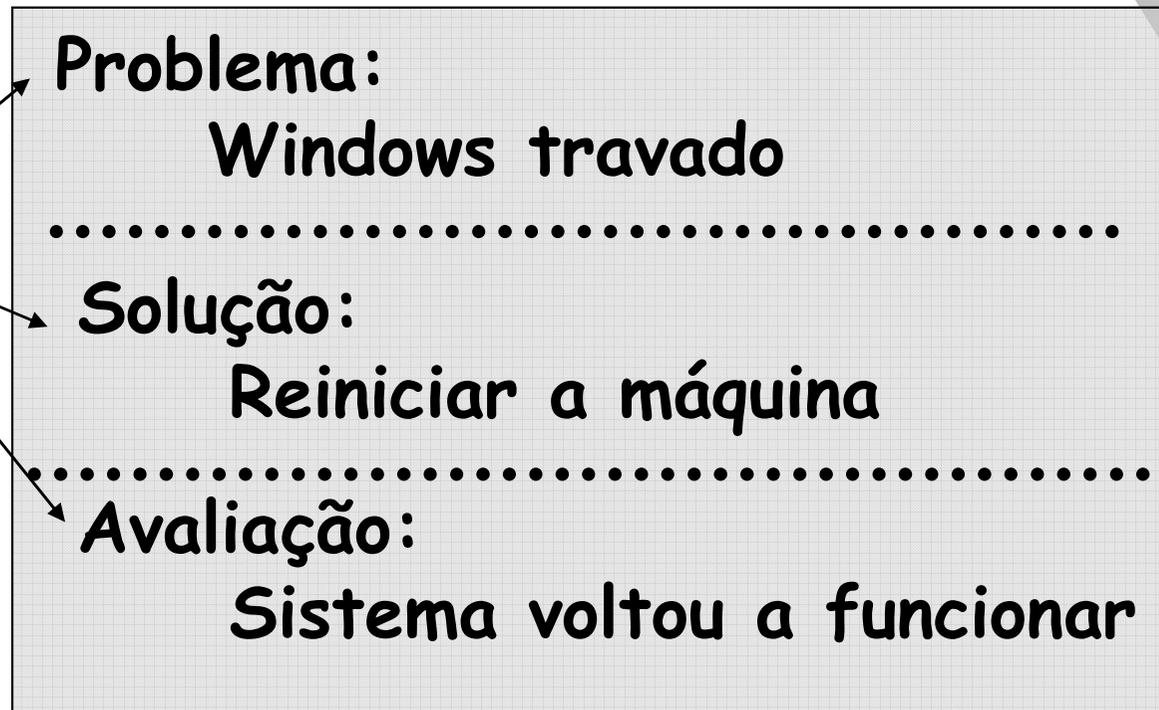
Componentes
de um caso



Exemplo de Caso

- ◆ Sistema para diagnóstico de falhas em computador
- ◆ Exemplo de caso :

Componentes
de um caso



Aplicações

- ♦ Diagnóstico
- ♦ Previsão
- ♦ Avaliação
- ♦ Planejamento
- ♦ Projeto
- ♦ Configuração



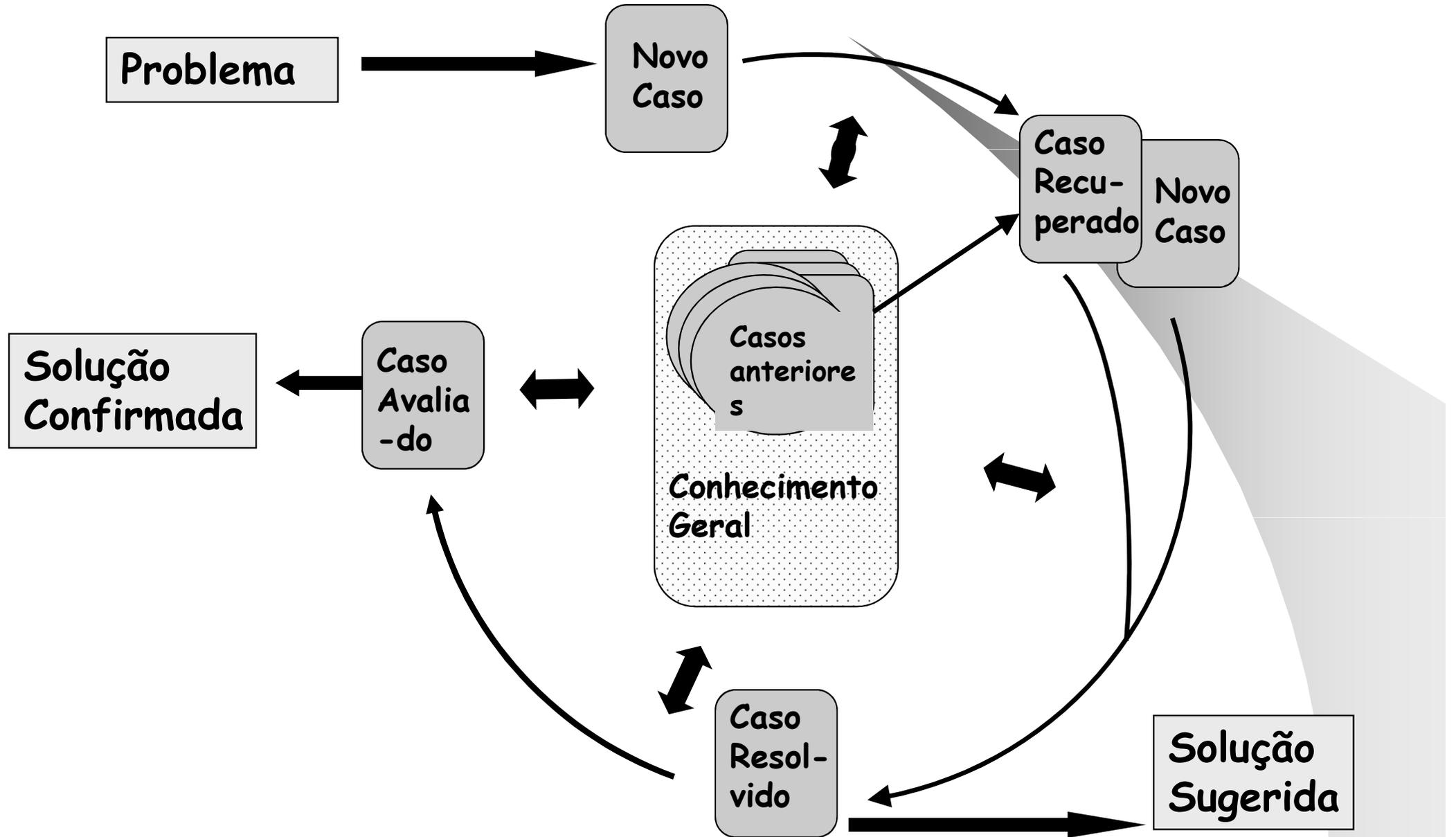
Aplicações

- ♦ *CYRUS* (Janet Kolodner, 1983)
 - sistema de perguntas e respostas sobre viagens e reuniões do Secretário de Estado Americano Cyrus Vance
- ♦ *PERSUADER* (Sycara, 1987)
 - soluciona conflitos entre patrões e empregados
- ♦ *CASEY* (Koton, 1989)
 - diagnostica problemas cardíacos
- ♦ *JULIA* (Hinrichs, 1992)
 - trabalha com planejamento de refeições
- ♦ *CHEF* (Hommond, 1996)
 - desenvolve novos pratos a partir de outros

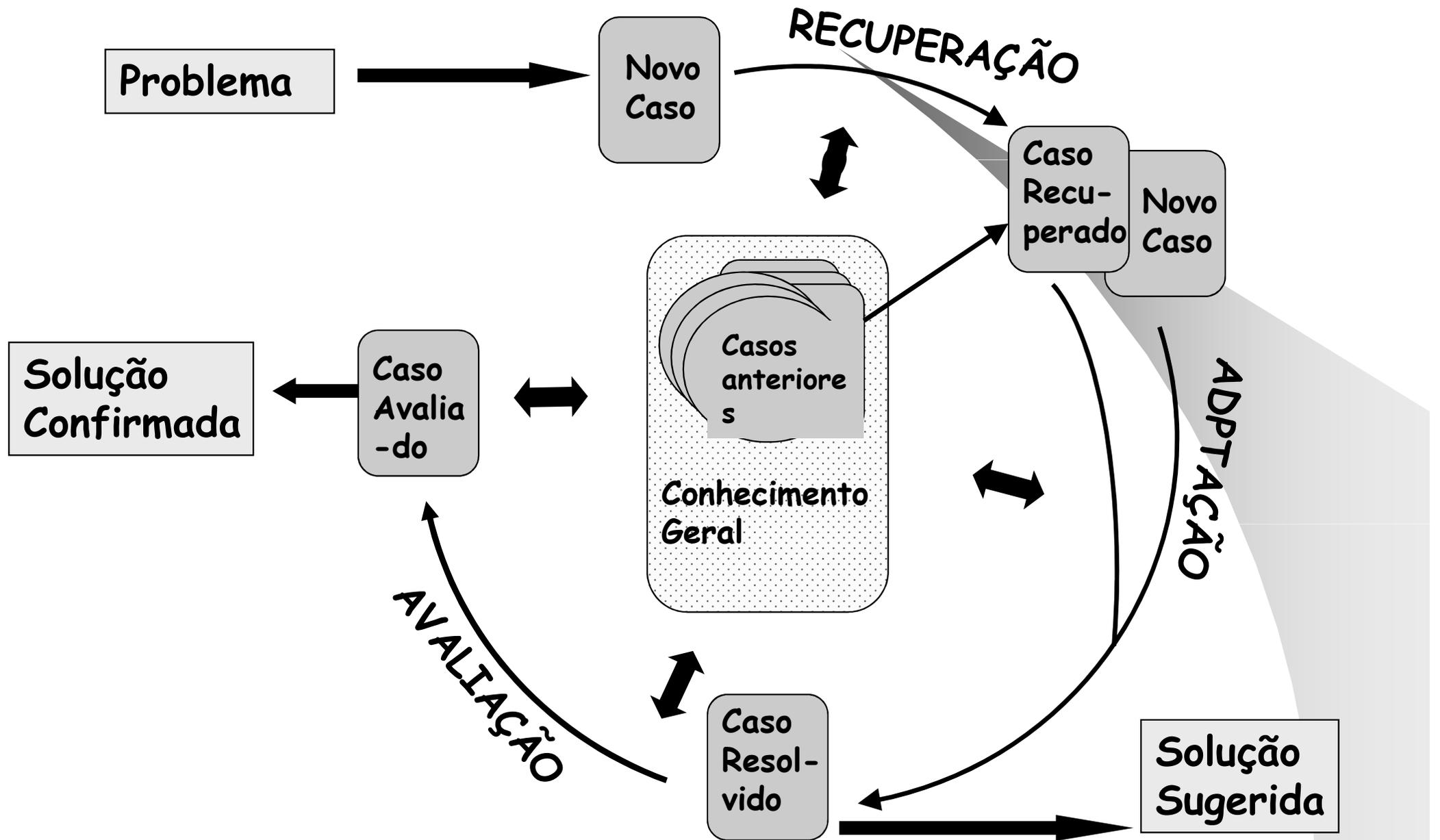
Aplicações

- ♦ *Sistemas de Assistência ao Cliente:*
 - Cisco Systems
 - Hewlett-Packard
 - Intel Corp
 - Microsoft
 - Visa International
 - AT&T Corp
 - Nokia Telecommunications

Funcionamento



Funcionamento



Representação e Organização de Casos

- ♦ A eficiência do sistema depende da estrutura e conteúdo da coleção de casos
- ♦ Problema de decidir :
 - O que armazenar em um caso (conteúdo)
 - Como estruturar seu conteúdo (estrutura)
 - Como organizar e indexar a memória de casos (organização e índice)

Representação de Casos

- ♦ Definir:
 - qual a estrutura adequada para os casos
 - quais casos devem ser representados
 - qual a granularidade da informação

Casos podem ser representados de várias formas, entre elas: frames, objetos, predicados,...

Objetos e tabelas do modelo relacional são as mais utilizadas.

Exemplo de caso

descrição do caso

- *Nome:* Paulo Rocha
 - *Nascimento:* 20.05.64
 - *Endereço:* Av. Carlos Gomes, POA
 - *Profissão:* Analista de sistemas
 - *Salário mensal:* R\$ 3.000,00
 - *Estado civil:* solteiro
 - *Dependentes:* 0
 - *Cartão crédito:* Visa
 - *Empréstimo solicitado:* R\$ 20.000,00
-

solução do caso

- *Empréstimo concedido:* sim
-

avaliação

- *cliente pagou corretamente o empréstimo*

Organização da Base de Casos

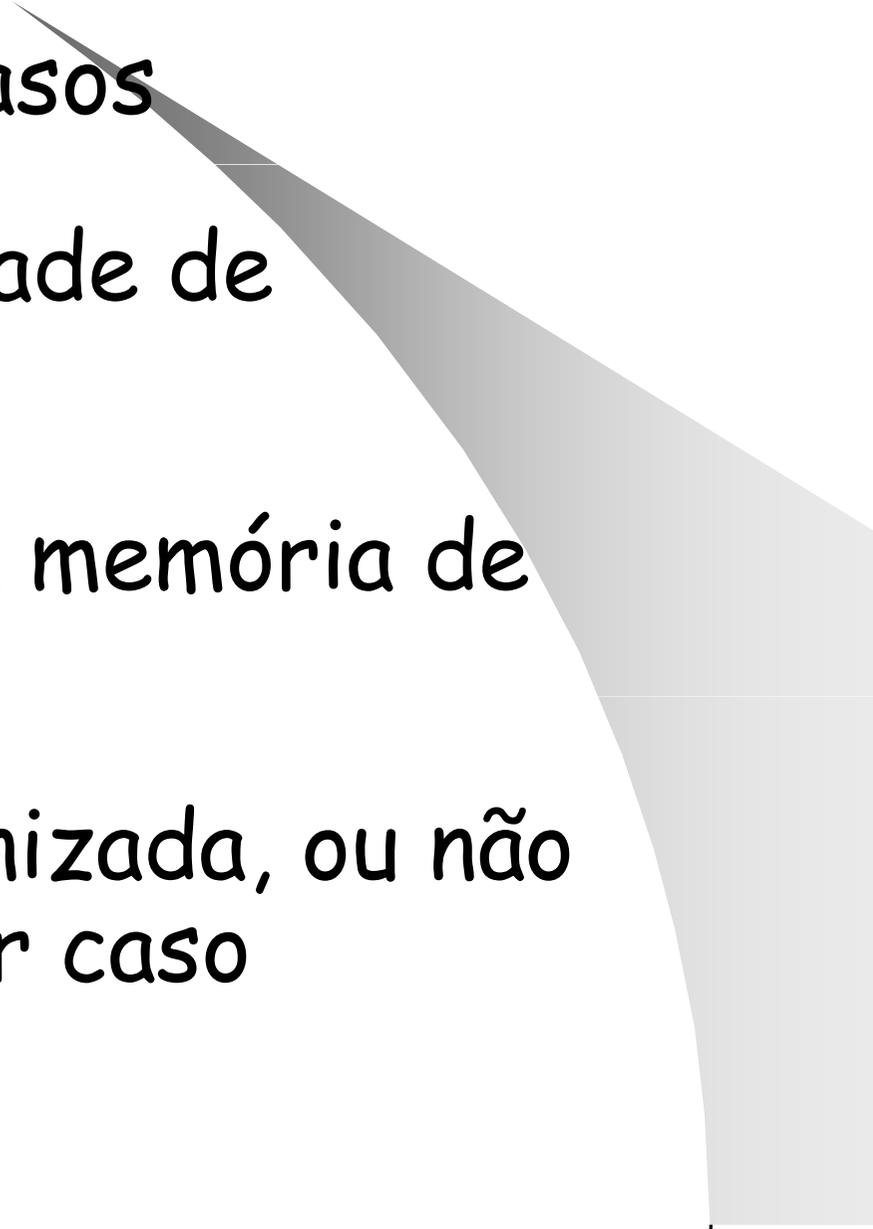
Influencia na recuperação do caso mais similar e nas atualizações da base de casos

- Organização Seqüencial
- Organização Estruturada

Organização Seqüencial

- ♦ Casos armazenados seqüencialmente em lista, *array* ou arquivo
- ♦ Ao fazer a recuperação, todos os casos são considerados
- ♦ Algoritmo simples para busca e atualização da base
- ♦ Ineficiente para bases muito grandes

Organização Estruturada

- ◆ Otimiza a busca de casos
 - ◆ Aumenta a complexidade de tratamento da base
 - ◆ Ocupa mais espaço na memória de trabalho
 - ◆ Deve estar bem organizada, ou não chegaremos ao melhor caso
- 

Organização da Base de Casos

João
Salário: 3000
Estado Civil: Solteiro
Dependentes: 0

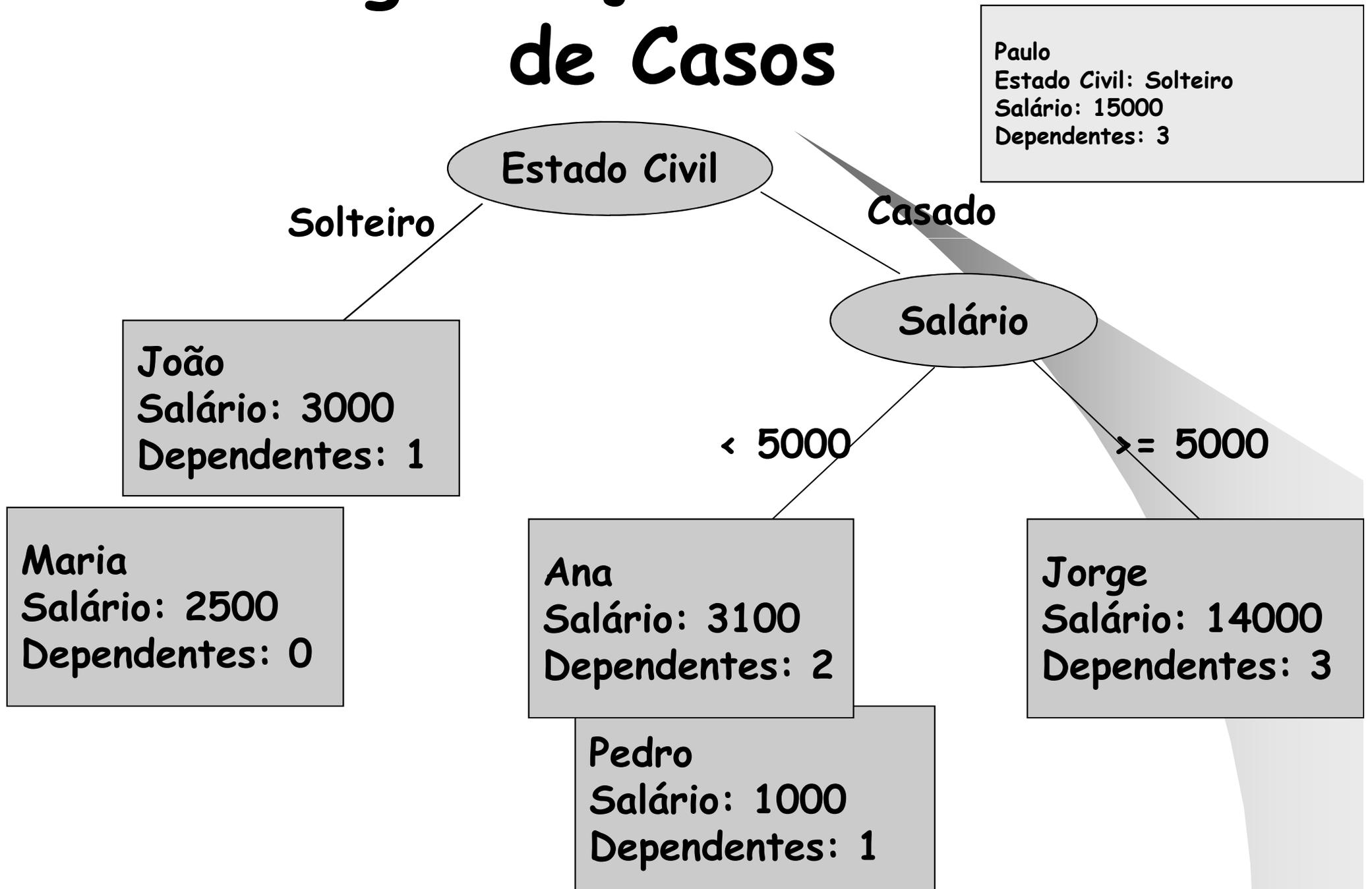
Maria
Salário: 2500
Estado Civil: Solteiro
Dependentes: 0

Ana
Salário: 3100
Estado Civil: Casada
Dependentes: 2

Jorge
Salário: 6000
Estado Civil: Casado
Dependentes: 3

Pedro
Salário: 10000
Estado Civil: Casado
Dependentes: 1

Organização da Base de Casos



Métodos de indexação

- ⇒ Indexar casos quer dizer definir caminhos (atalhos) que nos levam de fatos a casos.
- ⇒ Isto permite que quando estamos procurando pelo caso mais similar na base, não tenhamos que percorrer toda a base, mas possamos utilizar este "atalho"

- ♦ **As informações de um caso podem ser de dois tipos:**
 - **indexadas:** utilizadas na recuperação. Ex de diagnóstico médico: idade, sexo, tipo sanguíneo, peso
 - **não indexadas:** têm um valor de informação, mas não são usadas diretamente na recuperação. Ex: foto, endereço, nome do paciente, ...

Indexação

- ♦ **Técnicas Manuais**
 - Analisam caso a caso para determinar características que influenciam variações sobre as conclusões
- ♦ **Técnicas Automáticas**
 - Quantificam diferenças entre casos e relacionamentos entre características do problema e soluções adotadas

Métodos de indexação manuais

- ⇒ Manualmente a pessoa tem que analisar casos e dizer "este caso é importante por causa disto, ou daquilo" .
- ⇒ Uma das primeiras etapas na construção de um sistema com índices manualmente identificados é a definição de uma *checklist*;
- ⇒ Indexar desta forma é praticamente um trabalho de aquisição de conhecimento.

Indexação automática

- aprendizado indutivo: identifica as características que determinam as conclusões. Ex: ID3, C4.5
- indexação baseada em diferença: seleciona índices que diferenciam um caso de outro

Métodos de Recuperação

- ♦ Recuperar caso(s) mais similares
- ♦ Vários tipos de busca podem ser usadas:
 - serial, hierárquica, ...

Métodos de Recuperação

- ◆ Vizinho mais próximo (Nearest-Neighbour)

Para cada caso C_j da base

- Calcular a similaridade de C_j com o novo caso
- Reter o caso com o maior grau de similaridade

Vizinho mais próximo

$$\frac{\sum_{i=1}^n w_i \cdot \text{sim}(v^p_i, v^r_i)}{\sum_{i=1}^n w_i}$$

- ♦ W : peso da característica
- ♦ sim : função de similaridade
- ♦ v^p_i e v^r_i : valores da característica i

A função de similaridade depende do domínio do problema

Cálculo de Similaridade

Exemplo para tipo numérico:

$$a1 = 40$$

$$a2 = 80$$

$$\text{sim}(a1, a2) = 1 - |a2 - a1| / (\text{max} - \text{min})$$

Supondo que $\text{min} = 0$ e $\text{max} = 100$:

$$\text{sim}(40, 80) = 1 - |80 - 40| / (100 - 0) = 0,6$$

Cálculo de Similaridade

Exemplo para *strings*:

Cores = {Branco, Amarelo, Vermelho,
Marrom, Preto}

a1 = Branco

a2 = Amarelo

$$\text{Opção1: } \text{sim}(a1, a2) = \begin{cases} 1, & \text{se } a1 = a2 \\ 0, & \text{se } a1 \neq a2 \end{cases}$$

Cálculo de Similaridade

Opção2: enumerar distâncias uniformemente

Branco	Amarelo	Vermelho	Marrom	Preto
0	0,25	0,5	0,75	1

$a1 = \text{Branco}$

$a2 = \text{Amarelo}$

$$\text{sim}(a1, a2) = 1 - |0,25 - 0| / 1 = 0,75$$

Cálculo de Similaridade

Opção3: criar matriz de similaridades

	Branco	Amarelo	Vermelho	Marrom	Preto
Branco	1	0,8	0,4	0,15	0
Amarelo		1	0,5	0,2	0
Vermelho			1	0,7	0,6
Marrom				1	0,85
Preto					1

$a1 = \text{Branco}$

$a2 = \text{Amarelo}$

$\text{sim}(a1, a2) = 0,8$

Outro método de recuperação

Percorre estrutura de índice (ex: árvore de decisão) e no fim aplica o vizinho mais próximo para poucos registros