

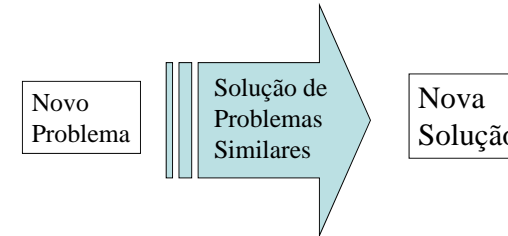
Raciocínio Baseado em Casos

Prof. Luis Otavio Alvares

prof. Luis Otavio Alvares

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

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



prof. Luis Otavio Alvares

CBR (*Case-Based Reasoning*)

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

Riesbeck & Schank, 1989

prof. Luis Otavio Alvares

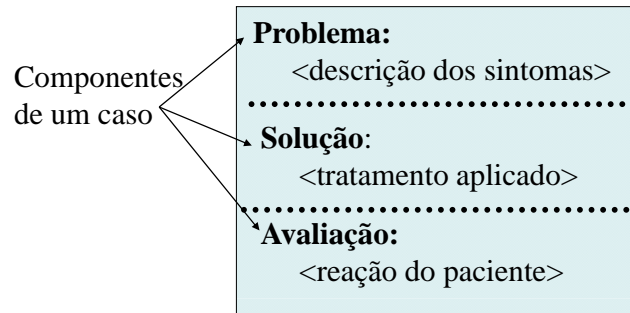
Caso

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

prof. Luis Otavio Alvares

Exemplo de Caso

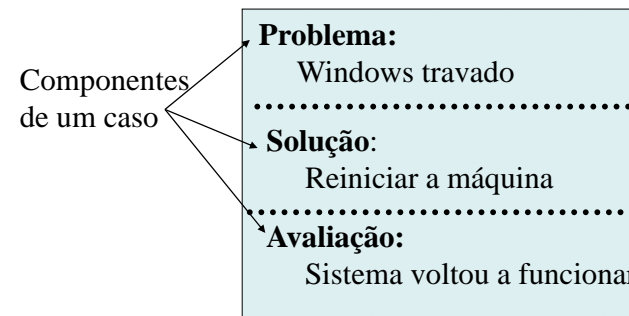
- Sistema para diagnóstico de doença
- Exemplo de caso :



prof. Luis Otavio Alvares

Exemplo de Caso

- Sistema para diagnóstico de falhas em computador - exemplo de caso :



prof. Luis Otavio Alvares

Aplicações

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

prof. Luis Otavio Alvares

Aplicações

- CYRUS (Janet Kolodner, 1983)
- 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*

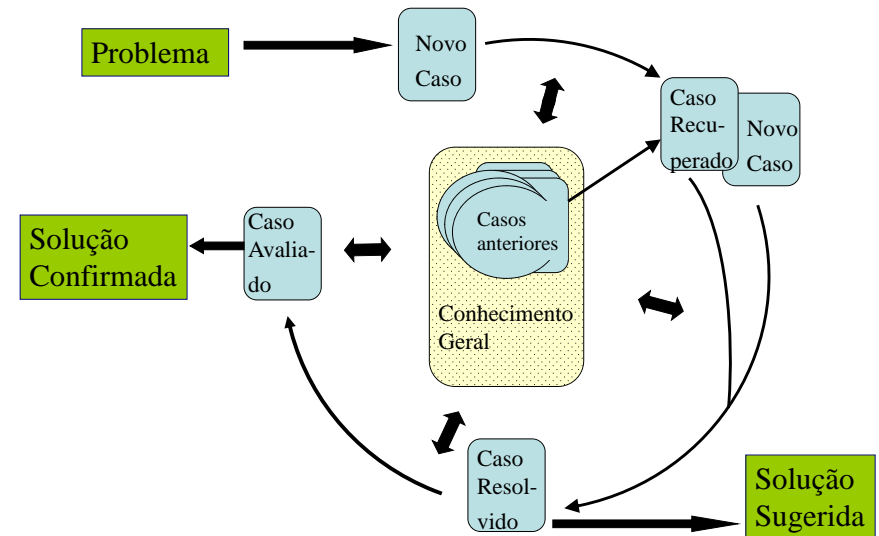
prof. Luis Otavio Alvares

Aplicações

- Sistemas de Assistência ao Cliente:
 - Cisco Systems
 - Hewlett-Packard
 - Intel Corp
 - Microsoft
 - Visa International
 - AT&T Corp
 - Nokia Telecommunications
 -
- Estimativa de custo de fabricação de peças na Peugeot

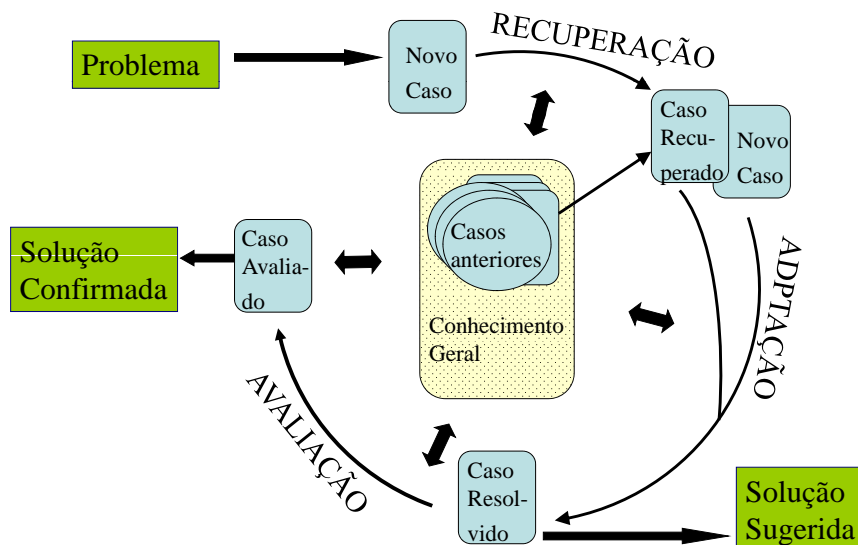
prof. Luis Otavio Alvares

Funcionamento



prof. Luis Otavio Alvares

Funcionamento



prof. Luis Otavio Alvares

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 o conteúdo (estrutura)
 - Como organizar e indexar a memória de casos (organização e índice)

prof. Luis Otavio Alvares

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.

prof. Luis Otavio Alvares

Exemplo de caso

descrição do caso

- Nome: Paulo Rocha
- Nascimento: 20.05.64
- Endereço: Av. Carlos Gomes, 12 - POA
- Profissão: Analista de sistemas
- Salário mensal: R\$ 5.000,00
- 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

prof. Luis Otavio Alvares

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

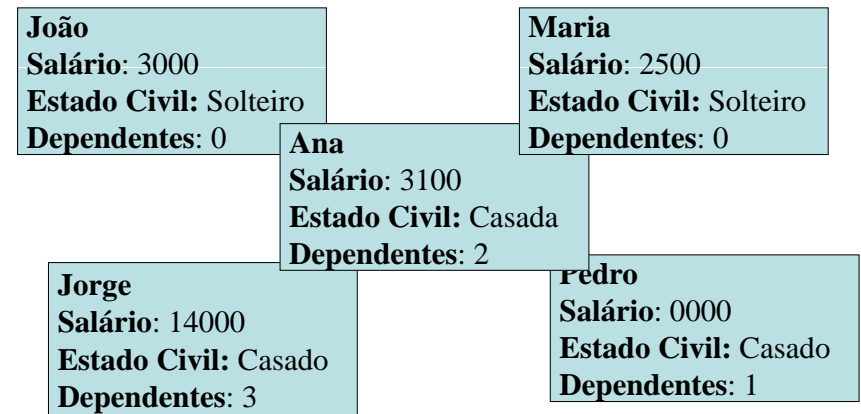
Organização Estruturada

- 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



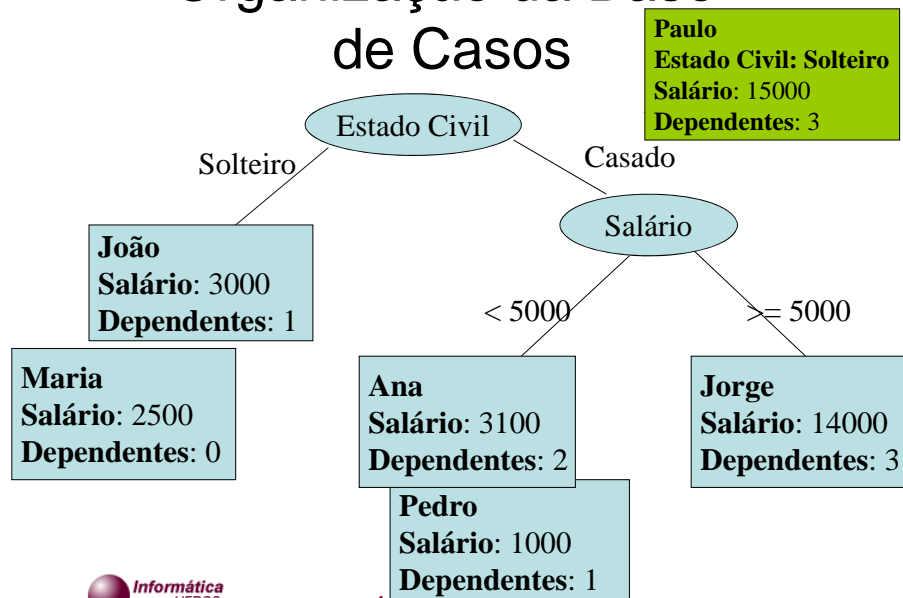
prof. Luis Otavio Alvares

Organização da Base de Casos



prof. Luis Otavio Alvares

Organização da Base de Casos



prof. Luis Otavio Alvares

Métodos de indexação

- ⇒ Indexar casos quer dizer definir caminhos (atalhos) que nos levam aos 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".



prof. Luis Otavio Alvares

Indexação

- 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, ...



prof. Luis Otavio Alvares

- 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



prof. Luis Otavio Alvares

Indexação manual

- ⇒ Manualmente a pessoa tem que analisar casos e dizer “as principais características desse caso são A e B...” .
- ⇒ Uma das primeiras etapas na construção de um sistema com índices manualmente identificados é a definição de um *checklist* de atributos (características) relevantes para o problema;
- ⇒ Indexar desta forma envolve um trabalho de aquisição de conhecimento.



prof. Luis Otavio Alvares

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



prof. Luis Otavio Alvares

Métodos de Recuperação

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



prof. Luis Otavio Alvares

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(s) caso(s) com o (k) maior(es) grau(s) de similaridade



prof. Luis Otavio Alvares

Vizinho mais próximo: cálculo da similaridade

$$\text{Sim}(c^n, c^r) = \frac{\sum_{i=1}^n w_i \cdot \text{sim}(v_i^n, v_i^r)}{\sum_{i=1}^n w_i}$$

- $\text{Sim}(c^n, c^r)$: similaridade entre os casos c^n e c^r
- w_i : peso da característica i
- $\text{sim}(v_i^n, v_i^r)$: função de similaridade considerando a característica i
- v_i^n e v_i^r : valores da característica i

A função de similaridade depende do domínio do problema



prof. Luis Otavio Alvares

Cálculo de Similaridade

Exemplo para tipo numérico:

$$a1 = 40 \quad \text{sim}(a1, a2) = ?$$

$$a2 = 80$$

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

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

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



prof. Luis Otavio Alvares

Cálculo de Similaridade

Exemplos para *strings*:

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

a1 = Branco

a2 = Amarelo

Opção1: medida binária de similaridade

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



prof. Luis Otavio Alvares

Cálculo de Similaridade

Opção2: enumerar distâncias uniformemente

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

a1 = Branco

a2 = Amarelo

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



prof. Luis Otavio Alvares

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 = Branco

a2 = Amarelo

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



prof. Luis Otavio Alvares

Outro método de recuperação

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



prof. Luis Otavio Alvares

Métodos de Adaptação

- É mais simples desenvolver regras de adaptação do que um sistema completo baseado em regras
- A maior parte dos sistemas comerciais apresenta o melhor caso e deixa a adaptação para o usuário



prof. Luis Otavio Alvares

Adaptação

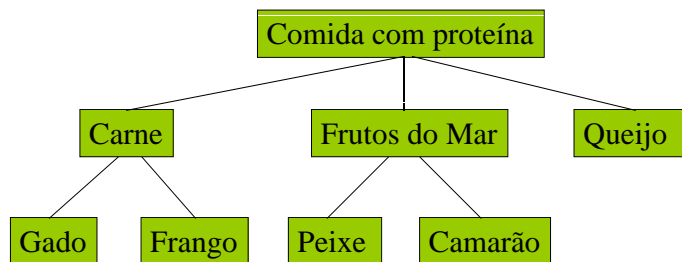
- Substituição
- Transformação



prof. Luis Otavio Alvares

Adaptação por substituição

- Instanciação de soluções antigas com valores substituíveis



Só é adequada se existe uma hierarquia ou outra forma de descrever o problema e que nos permita substituir um elemento por outro



prof. Luis Otavio Alvares

Transformação

- Funciona através de heurísticas
- Exemplo: lasanha vegetariana
 - remover ingrediente secundário (ex: brócolis)
 - substituir item (ex: colocar beringela)



prof. Luis Otavio Alvares

Ferramentas

- AcknoSoft - KATE
- Brightware Inc. - ART*Enterprise
- Case Bank Support Systems Inc. - Spotlight
- Cognitive Systems Inc. - ReMind
- Esteem Software Inc. - ESTEEM
- Inductive Solutions Inc. - CasePower



prof. Luis Otavio Alvares

Ferramentas

- Inference Corporation - k-commerce (CBR3 or CBR Express, CasePoint, Generator & WebServer)
- IET-Intelligent Electronics - TechMate
- Intellix - KnowMan
- Isoft - ReCall
- Sententia Software Inc. - CASE Advisor & Case Advisor Webserver
- ServiceSoft - Knowledge Builder & Web Adviser



prof. Luis Otavio Alvares

Softwares Acadêmicos

- CASPIAN (University of Wales)
- CASUEL - A Common Case Representation Language
- PROTOS
- EProtos - extensão do Protos



prof. Luis Otavio Alvares

Comparação com um Banco de Dados tradicional

- BD não contém a descrição dos problemas associados com a solução
- BD não faz casamento por aproximação
- a base de casos não é livre de ruídos e duplicações



prof. Luis Otavio Alvares

Vantagens do CRB

- Construção de um protótipo antes de obter a completa estruturação do domínio
- Reuso de conhecimento armazenado em bancos de dados e outras fontes
- Diminui a necessidade de aquisição de conhecimento
- Aprendizagem “automática” de novos casos



prof. Luis Otavio Alvares

Desvantagens do CBR

- Dificuldade em obter casos disponíveis e confiáveis
- Não cobrem todo o domínio
- Não existem bons algoritmos de adaptação
- Exige mais espaço para armazenamento



prof. Luis Otavio Alvares

Referências

Kolodner, Janet. **Case-Based Reasoning**. San Mateo, California: Morgan Kaufman, 1993.

Aamodt, A. & Plaza, E. **Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches**. *AICom - Artificial Intelligence Communications, IOS Press, Vol.7: 1, pp.39-59, 1994.*

Watson, Ian & Marir, Farhi. **Case-Based Reasoning: A Review**. *AI_CBR, University of Salford, Salford, M5 4WT, UK*

[<http://www.ai-cbr.org/>]



prof. Luis Otavio Alvares

Exemplo de software:
- CASPIAN -

prof. Luis Otavio Alvares

CASPIAN

- Shell publicamente disponível na University of Wales:
http://www.aber.ac.uk/compsci/Research/mbsg/cbrprojects/getting_caspian.shtml
- CASPIAN cria uma base de casos a partir de arquivo de casos escrito em CASL
- comando: `caspian <nome_do_arquivo>`

Funcionamento do CASPIAN

- ◆ Representação dos Casos
 - utiliza a linguagem CASL
- ◆ Indexação
- ◆ Recuperação
- ◆ Adaptação

CASL

- Linguagem usada para descrever casos
- Casos são descritos em um arquivo tipo texto
- O arquivo de casos é usado pelo CASPIAN para criar a base de casos no computador

Sintaxe

[Introdução]
Definição do Caso
Definição do Índice
[Definição da Modificação]
[Definição da regra de adaptação]
Instância do Caso {}

Sintaxe de um caso

[Introdução]
Definição do Caso
Definição do Índice
[Definição da Modificação]
[Definição da regra de adaptação]
Instância do Caso {}

Ajuda o usuário a entender o propósito da base de casos

Seção Introdução (opcional)

```
introduction is
'Base de Casos CHEF,
versao modificada'
end;
```

- Strings são contidas entre aspas simples

Sintaxe de um caso

[Introdução]
Definição do Caso
Definição do Índice
[Definição da Modificação]
[Definição da regra de adaptação]
Instância do Caso {}

Define tipos e pesos dos atributos do problema.

É usado para validar os casos e a entrada do usuário

Definição do Caso

```
CASE DEFINITION IS
  FIELD sabor TYPE IS (doce, salgado)
  WEIGHT IS 3
  PROMPT IS ['Entre o sabor do prato'];
  FIELD ...
SOLUTION DEFINITION IS
  FIELD ...
END;
```

- Os atributos *weight* e *prompt* são opcionais
- São os campos que pertencem a definição do problema
- O atributo *weight* é usado no processo de casamento

Sintaxe de um caso

[Introdução]
Definição do Caso
Definição do Índice
[Definição da Modificação]
[Definição da regra de adaptação]
Instância do Caso {}

Define os campos que serão usados como índices.

Definição do Índice (obrigatório)

```
INDEX DEFINITION IS  
  INDEX ON sabor  
END;
```

- Um campo deve ser usado como índice somente se for muito importante
- Usado pelo CASPIAN para gerar a estrutura de índices

Sintaxe de um caso

[Introdução]
Definição de Caso
Definição de Índice
[Definição de Modificação]
[Definição de regra de adaptação]
Instância do Caso {}

Possui lista de valores similares, para guiar o processo de casamento

Definição de Modificação

```
MODIFICATION DEFINITION IS  
  ABSTRACTION carne_vermelha IS (vaca, carneiro, porco);  
  ABSTRACTION qualquer_carne IS (carne_vermelha,  
  carne_branca);  
  FIELD tempo_preparo SIMILAR RANGE 15 TO 30;  
  FIELD tempo_preparo SIMILAR RANGE 25 TO 35;  
  FIELD tempo_preparo SIMILAR RANGE 45 TO 60;  
END;
```

Sintaxe de um caso

[Introdução]
Definição do Caso
Definição do Índice
[Definição da Modificação]
[Definição da regra de adaptação]
Instância do Caso { }

Usadas para modificar a solução recuperada do banco de dados

Definição de Regras de Adaptação

```
REPAIR RULE DEFINITION IS
REPAIR RULE prepare_legume IS
  WHEN
    vegetal is legume
  THEN
    CHANGE preparacao_vegetal TO descascar;
  END;
END;
```

Sintaxe das Regras de Adaptação

- A condição *when* pode ser composta:
 - or, and, not
- Atributos podem ser comparados com:
 - =, !=, <, >, <=, >=
- Pode-se utilizar expressões aritméticas:
 - +, -, *, /, ^
- Outras ações que podem ser tomadas:
 - **change** <identificador> **to undefined**;
 - **change weight of** <identificador> **to** <numero>
 - **reselect**;
 - **repair**;

Sintaxe de um caso

[Introdução]
Definição do Caso
Definição do Índice
[Definição de Modificação]
[Definição de regra de adaptação]
Instância do Caso { }

Casos que fazem parte da base

Definição de um caso

```
case instance mange_tout is
  ingredients = [ green_beans ];
  cook_method = stir_fry;
  taste = savoury;
  vegetable = green_beans;
solution is
  recipe = [ cook_method the ingredients ];
end;
```

Recuperação no CASPIAN

A busca é realizada baseada no valor dos índices

- resultado da busca é um conjunto de casos candidatos

Recuperação no CASPIAN

Para selecionar o melhor caso:

- CASPIAN usa o algoritmo do vizinho mais próximo, ou
- Usuário percorre a lista de casos candidatos e seleciona o melhor

Adaptação de Casos

- Valores originais do caso recuperado são alterados pelos valores do usuário
- Regras de adaptação são aplicadas