



20 e 21 de outubro  
Instituto Nacional de Pesquisas Espaciais - INPE  
São José dos Campos - SP

## Uma Ferramenta de Agrupamento para Raciocínio Baseados em Casos

Marilyn Menecucci Ibañez<sup>1</sup>, Sandra Sandri<sup>2</sup>

<sup>1</sup>Programa de Mestrado ou Doutorado (conforme cada aluno) em Computação Aplicada – CAP  
Instituto Nacional de Pesquisas Espaciais – INPE

<sup>2</sup>Laboratório Associado de Computação e Matemática Aplicada – LAC  
Instituto Nacional de Pesquisas Espaciais – INPE

marilynmenecucciibañez@gmail.com<sup>1</sup>, sandri@lac.inpe.br<sup>2</sup>

**Abstract.** *The objective of this project is to develop a tool for applying the technique of similarity between cases (CRR - Case Resemblance Relation) in a case base. The technique involves the creation of groups (clusters) of similar cases by calculating the similarities of the cases the base. This article is used as example cases in the data of cars.*

**Resumo.** *O objetivo deste projeto é desenvolver uma ferramenta que aplique a técnica relação de semelhança entre casos (CRR - Case Resemblance Relation) em uma base de casos. A técnica consiste na criação de agrupamentos (clusters) de casos similares por meio do cálculo das similaridades dos casos da base. Neste artigo utiliza-se como exemplo ilustrativo casos referentes a dados de carros.*

**Palavras-chave:** *Raciocínio Baseado em Casos, Relações de Similaridades, Sistemas Difusos*

### 1. Introdução

As pessoas costumam utilizar de forma natural situações vividas em seu cotidiano para a resolução de novos problemas. Isto pode ser feito selecionando-se as situações passadas mais semelhantes ao novo problema para que se possa ter uma solução final. A idéia de semelhança também é bastante aplicada em diversas áreas de pesquisas tais como na pesquisa da prevalência da esquistossomose e monitoramento de florestas. Assim, com base neste contexto este trabalho utiliza como técnica base a Raciocínio Baseado em Casos (CBR - Case-Based Reasoning) que tem sido aplicada com êxito em diversas áreas do conhecimento, nas quais existem ao mesmo tempo um grande volume de exemplos e alguma maneira de se medir as semelhanças entre esses exemplos. Desta forma, o trabalho aborda uma técnica de CBR, recentemente proposta, que se baseia numa relação de semelhança entre casos (CRR - Case Resemblance Relation). Segundo essa técnica,

dois casos são semelhantes quando a similaridade entre as soluções de seus problemas é tão grande quanto, ou superior, à similaridade (não nula) entre a descrições dos problemas em si. A relação de similaridade entre casos permite a criação de agrupamentos (clusters) de casos similares. Ao se apresentar um novo problema à base de casos, verifica-se o quanto esse problema é similar a cada agrupamento e uma solução para o problema é calculada a partir das soluções dos casos no agrupamento. As similaridades e soluções relativas aos diversos agrupamentos tomados individualmente são então agregados para produzir uma solução final para o novo problema. A criação dos agrupamentos requer um grande esforço computacional. Nesse trabalho apresenta-se uma ferramenta para a criação e manipulação dos agrupamentos que representa um primeiro passo na direção de de um objetivo maior, o de minimizar o custo da criação de agrupamentos em bases com grande número de casos. Neste trabalho, ilustra-se o desenvolvimento e uso da ferramenta em uma base de casos criada como ilustração.

## 2. Metodologia

### 2.1. Raciocínio Baseado em Casos

A técnica CBR do inglês *Case-Based Reasoning* ou Raciocínio Baseado em Casos consiste em utilizar um conjunto de soluções anteriores com ou sem alterações, dentro de um determinado domínio, para solucionar novos problemas [2]. De acordo com [1], o CBR, de uma forma generalizada, pode ser dividido nas seguintes etapas.

- Recuperação dos casos mais semelhantes.
- Reutilização da informação e conhecimento do caso para resolver o problema.
- Revisão da solução proposta, podendo ser adaptada a uma nova realidade.
- Retenção de partes da experiência obtida para ser utilizadas na resolução de novos problemas.

A Figura 1 [6] apresenta o ciclo dos 4 R's do RBC. Esse ciclo mostra como essas etapas estão relacionadas em tal processo.

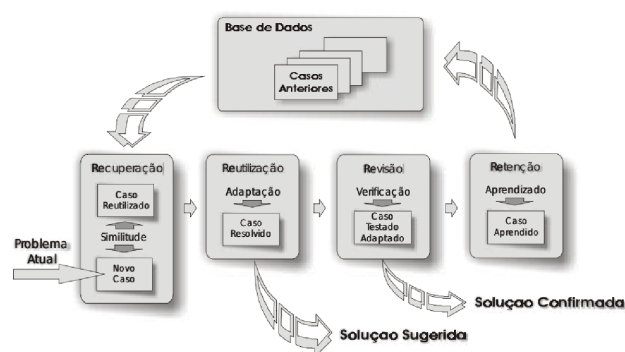


Figura 1. Ciclo dos 4 R's do RBC.

### 2.2. Técnica Utilizada: Relação de Semelhança de Casos - CRR

Seja  $v$  uma variável definido em um domínio  $\Omega$ .  $S_v : \Omega \rightarrow [0, 1]$  é uma relação de similaridade se  $S_v$  for simétrica e reflexiva. Uma relação similaridade nebulosa  $S$  em um domínio  $\Omega$  é um mapeamento  $S : \Omega \times \Omega \rightarrow [0, 1]$ , que atribui a cada par  $(\omega, \omega')$  de elementos de  $\Omega$  um valor que mede o quanto  $\omega$  and  $\omega'$  são similares. A relação de similaridade obedece usualmente as seguintes condições:

- Reflexividade:  $S(\omega, \omega) = 1$
- Simetria:  $S(\omega, \omega') = S(\omega', \omega)$

Alguns autores incluem também a transitividade:

- Transitividade:  $S(\omega, \omega'') \geq S(\omega, \omega') \otimes S(\omega', \omega'')$ , onde  $\otimes$  é uma t-norma

Em particular,  $S(\omega, \omega') = 1$  significa que  $\omega$  e  $\omega'$  são indistinguíveis, enquanto  $S(\omega, \omega') = 0$  significa que  $\omega$  e  $\omega'$  não possuem nada em comum. Pode-se também entender  $S(\omega, \omega') = 1 - S(\omega, \omega')$  como um tipo de distância entre  $w$  e  $w'$ . Neste trabalho apenas considera-se as relações que satisfaçam reflexividade e simetria [4].

No restante deste trabalho, será utilizada a seguinte família parametrizada de relações de similaridade:

$$\forall a, b \in \Omega, S_v(a, b) = \max(0, 1 - \frac{|a - b|}{\lambda})$$

onde  $\lambda > 0$ .

O valor extremo  $S_v(a, b) = 1$  (respec.  $S_v(a, b) = 0$ ) indica que  $a$  e  $b$  são considerados completamente similares (respec. dissimilares). Valores de similaridade entre 0 e 1 indicam similaridade parcial.

A *Case Resemblance Relation* ou Relação de Semelhança de Casos  $S_{res}$  é derivada a partir de duas relações:  $S_{in}$ , que descreve a similaridade dos casos no espaço do problema e  $S_{out}$  que descreve a similaridade dos casos no espaço da solução.

Seja  $V = (v_1, \dots, v_n)$  um conjunto de variáveis de descrição de problemas, onde cada  $v_i$  está definido em um domínio  $\Omega_i$ . Seja  $p = (p_1, \dots, p_n)$  e  $p' = (p'_1, \dots, p'_n)$  dois problemas, onde  $p_i, p'_i \in \Omega_i$ . A cada variável  $v_i$  está associada uma relação de similaridade  $S_{v_i}$ , como descrita acima. A relação  $S_{in}$  pode ser derivada a partir das relações  $S_{v_i}$  de várias maneiras. Neste trabalho, obtém-se  $S_{in}$  pela média aritmética:

$$S_{in}(p, p') = \sum_{k=1, n} \frac{1}{n} \times S_{v_i}(p_i, p'_i)$$

Neste trabalho, considera-se que a base de casos concerne somente uma variável de solução  $o$ . A relação  $S_{out}$  é obtida diretamente de  $S_o$ , i.e.  $S_{out} = S_o$ .

Seja  $B = \{c_j = (d_j, o_j)\}_{j=1, m}$  uma base de casos, onde cada descrição de problema  $d_j$  é uma n-upla em  $\Omega_1 \times \dots \times \Omega_n$ .

Para o cálculo do  $S_{res}$  utiliza-se os valores calculados para  $S_{in}$  e  $S_{out}$  da seguinte maneira:

$$\forall c_k, c_l \in B, S_{res}(c_k, c_l) = \text{1if} 0 < S_{in}(d_k, d_l) \leq S_{out}(o_k, o_l)$$

A CRR é a base para a construção de um processo de semelhança de hipergrafos (CHR), no qual os *hyperedges* são *clusters* dos casos relacionados com o CRR.

Quando uma nova descrição é apresentada, compara-se com os conjuntos de casos semelhantes na base de casos, seleciona-se o melhor caso e obtém-se a nova solução a partir das soluções dos casos no grupo selecionado. Portanto, nossa abordagem pode ser dividida basicamente em três tarefas principais:

- criação de conjuntos de (eventualmente se cruzam) clusters de casos semelhantes na base de caso.
- seleção do cluster cujos casos de descrições atingem o grau de similaridade mais próximos com a descrição do novo caso,
- cálculo da solução para o problema de novo, em função das soluções geradas por casos individuais no conjunto selecionado.

Finalmente, uma quarta tarefa poderia ser prevista em que a solução de vários agrupamentos seriam agregadas [4].

### 2.3. Identificação dos Clusters

Seja  $B$  uma base com  $m$  casos e  $P(B)$  o conjunto potência de  $B$ . O conjunto de *clusters*  $H(B) \in P(B)$  é identificado utilizando o resultado encontrado pelo cálculo da relação  $S_{res}$ :  $\forall c_1, c_2 \in B$ , se  $\exists h \in H(B)$  tal que  $c_1, c_2 \in h$ , então  $S_{res}(c_1, c_2) = 1$ .

Nesta relação, para cada caso similar ao outro, ou seja, que o valor da relação seja igual a 1, é criado um *cluster*. Se um dos casos da relação já estiver contido em outro *cluster*, verifica-se se o caso não pertencente possui relação de similaridade igual a 1 com os demais casos do *cluster*. Se similar o caso é adicionado no *cluster* existente, no contrário cria-se um novo *cluster* com os casos similares.

Neste trabalho, são somente considerados *clusters* maximais. Um cluster  $A \in H(B)$  é máximo se  $\nexists C \subseteq H(B), A \subset C$ .

### 3. Determinação da Força dos Clusters (Cluster Strength)

Em um sistema CBR, a seleção de cluster pode ser visto como parte do processo de recuperação. O cluster cujas descrições atingir o grau de similaridade geral com a descrição do novo caso é selecionado. Para obter o máximo de similaridade geral vamos calcular o que chamamos de força de um cluster (*cluster strength*)  $A = \{c_1, \dots, c_n\}, c_i = (p_i, s_i)$  em relação a um problema  $p^*$  do acordo com a função  $f$  [4], dada por:

$$str(A, p^*) = f(S_{in}(p_1, \dots, p^*), \dots, S_{in}(p_n, \dots, p^*)).$$

Exemplos de operadores de (*cluster strength*) são apresentados a seguir:

- T-norma ( $str_X$ ) (produto):

$$\forall a, b \in [0, 1], \times(a, b) = a \times b$$

- T-conorma ( $str_{bs}$ ) soma limitada (bounded sum):

$$\forall a, b \in [0, 1], bs(a, b) = (a + b - ab)$$

- Média (operadores monotônicos, usualmente não associativos, entre o máximo e mínimo)  $str_m$  (média aritmética):

$$\forall a, b \in [0, 1], m(a, b) = \frac{(a + b)}{2}$$

#### 4. Computação de solução para novo problema

Computar uma solução única a partir de um cluster (conjunto de clusters) pode ser visto como parte do processo de reuso ou da Fase de Validação. Calcula-se a solução para um problema  $p^*$  em relação a um conjunto  $h = \{c_1, \dots, c_n\}$ ,  $c_i = (p_i, s_i)$ , de acordo com uma função  $g$ , dada por:

$$sol_g(p^*, h) = g(s_1, \dots, S_n)$$

Até o momento, a média ponderada tem sido usada como a função de agregação  $g$ . Os pesos são as similaridades da descrição dos casos com o novo problema  $p^*$ :

$$g(p^*, h) = \frac{\sum_{c_i=(p_i, o_i) \in h} (S_{in}(p_i, p^*) \times o_i)}{\sum_{c_i=(p_i, o_i) \in h} S_{in}(p_i, p^*)}$$

### 5. Sistema Computacional

O sistema computacional consiste de duas fases a primeira fase esta relacionada ao pré-processamento dos dados de entrada gerando um modelo e a segunda fase ao processamento do modelo criado.

As próximas seções apresentam a modelagem e os algoritmos desenvolvidos para a primeira fase do sistema.

#### 5.1. Primeira Fase

Na primeira fase do sistema computacional realiza-se um pré-processamento dos dados da base de dados escolhida para ilustrar a funcionamento do programa.

##### 5.1.1. Modelagem

A modelagem do programa se baseia nas visões e diagramas padrões da UML 2.0. A UML 2.0 é um conjunto de diagramas que ajuda na descrição e desenvolvimento de um sistema de software [5]. Cada diagrama possui características próprias que ajudam a definir melhor as propriedades do projeto a ser desenvolvido. As visões da UML constituem uma projeção na organização e estrutura do sistema. Existem cinco visões necessárias para o desenvolvimento e implantação do sistema.

**Caso e Uso:** descreve a funcionalidade do sistema desempenhada pelos atores externos do sistema (usuário).

**Lógica:** descreve como a funcionalidade do sistema será implementada.

**Componentes:** descrição da implementação dos módulos e suas dependências.

**Concorrência:** apresenta a divisão do sistema em processos e processadores.

**Organização:** mostra a organização dos computadores e periféricos e como eles se conectam entre si.

## **Análise de Requisitos**

Os requisitos necessários para o início da construção do software são listados a seguir:

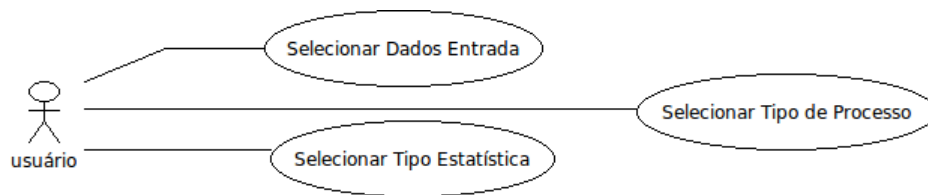
- Entrada de dados: Tabela com a base de casos e valores de lambdas ou tabelas com o cálculo das similaridades.
- Cálculo e/ou apresentação dos valores das Similaridades dos casos.
- Cálculo dos valores de entrada  $S_{in}$
- Cálculo dos valores de saída  $S_{out}$
- Cálculo de  $S_{res}$
- Cálculo dos *Clusters*
- Cálculo das estatísticas dos *Clusters*

## **Diagramas UML**

A seguir é apresentado dois modelos de diagramas UML para representar o sistema desenvolvido.

### **Diagrama de Casos de Uso**

O diagrama de casos de uso especifica uma sequência de ações, inclusive variantes, que um sistema realiza e que produz um observável resultado de valor para um particular ator [5]. A Figura 2 mostra este diagrama aplicado no sistema.



**Figura 2. Diagrama de casos de uso do programa.**

### **Diagrama de Classes**

Um diagrama de classes descreve os tipos de objetos no sistema e os vários tipos de relacionamento estático existente entre eles, bem como atributos e operações de uma classe e as restrições [5]. A Figura 3 mostra este diagrama aplicado no sistema.

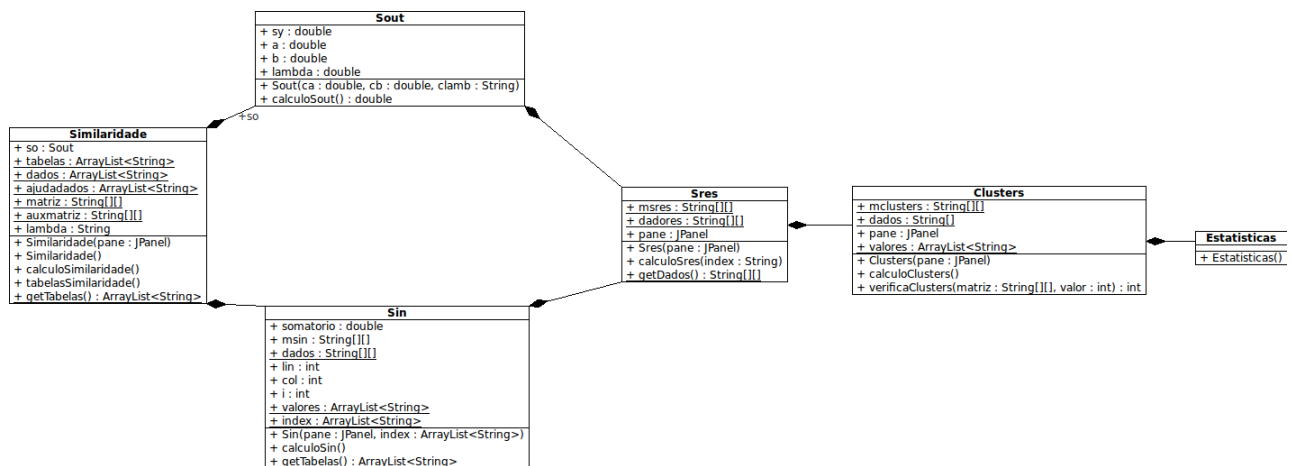


Figura 3. Diagrama de classe do programa.

### 5.1.2. Telas Iniciais Programa

O programa desenvolvido na linguagem Java [3], utiliza os algoritmos Sin, Sout e Sres, Clusters, Produto, Bounded Sum e Media. O software é constituído pelos seguintes itens.

- **menu Arquivo**

**Abrir:** selecionar o arquivo com os dados do case base.

**Sair:** finaliza a execução e sai do programa.

- **menu Processar**

**Similaridade:** apresenta duas opções de entrada para a similaridade - Escolha do arquivo com as tabelas de similaridades já calculadas ou escolha do arquivo com valores de  $\lambda$  para o cálculo da similaridade.

**Sin:** cálculo dos valores de entrada utilizando as similaridades dos dados da base de casos.

**Sout:** cálculo do valor de saída de uma determinada variável utilizando os dados da base de casos.

**Sres:** resultado final do cálculo da similaridade utilizando os dados de *Sin* e *Sout*.

**Clusters:** apresentação dos clusters encontrados por meio do resultado final *Sres*.

- **menu Estatística:**

**Produto:** apresenta o cálculo do produto da similaridade dos casos pertencentes a um determinado cluster.

**Bounded Sum:** calcula uma soma específica da similaridade dos casos pertencentes a um determinado cluster.

**Media:** realiza o cálculo da média aritmética da similaridade dos casos pertencentes a um determinado cluster.

A Figura 4 apresenta o menu de seleção da forma de apresentar/calcular a *Similaridade*, a Figura 5 apresenta o menu de seleção dos atributos para o cálculo de  $S_{in}$  e a Figura 6 apresenta o menu das opções para o cálculo das estatísticas dos *Clusters*.



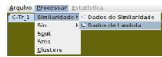


Figura 4. Detalhe do menu de seleção da forma de apresentar/calcular a *Similaridade*.



Figura 5. Detalhe do menu dos atributos para o cálculo de  $S_{in}$ .



Figura 6. Detalhe do menu das opções para o cálculo das estatísticas dos *Clusters*.

## 5.2. Resultados

O objetivo deste capítulo é apresentar os resultados desenvolvidos até o presente momento do sistema computacional para a realização do cálculo da similaridade de um conjunto de dados e sua aplicação na implementação do programa para tal cálculo.

### 5.2.1. Base de Casos

A Figura 7 apresenta a base de casos utilizada para teste do programa desenvolvido.

	cylinder	displacement	hp	weight	acceleration	year	origin	mpg
C1	8	307	130	3504	12	70	1	18
C2	4	97	88	2130	14,5	71	3	27
C3	3	70	97	2330	13,5	72	3	19
C4	6	225	105	3121	16,5	73	1	18
C5	4	98	83	2219	16,5	74	2	29
C6	8	262	110	3221	13,5	75	1	20
C7	4	90	70	1937	14,2	76	2	29
C8	6	250	110	3520	16,4	77	1	17,5
C9	5	131	103	2830	15,9	78	2	20,3
C10	5	183	77	3530	20,1	79	2	25,4

Figura 7. Apresentação da base de casos.

### 5.2.2. Similaridade dos dados da Base de Casos

As Figuras 8, 9, 10, 11, 12, 13, 14, 15 apresentam os resultados de *Sout* para *cyl*, *displ*, *hp*, *weight*, *accel*, *year*, *orig* e *mpg*.

cyl	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.0	0.5	0.0	1.0	0.0	0.5
c2	0.0	1.0	0.75	0.5	1.0	0.0	1.0	0.5
c3	0.0	0.75	1.0	0.25	0.75	0.0	0.75	0.25
c4	0.5	0.5	0.25	1.0	0.5	0.5	0.5	1.0
c5	0.0	1.0	0.75	0.5	1.0	0.0	1.0	0.5
c6	1.0	0.0	0.0	0.5	0.0	1.0	0.0	0.5
c7	0.0	1.0	0.75	0.5	1.0	0.0	1.0	0.5
c8	0.5	0.5	0.25	1.0	0.5	0.5	0.5	1.0

Figura 8. Relação de Similaridade para *cyl*.

displ	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
c2	0.0	1.0	0.0	0.0	0.95	0.0	0.65	0.0
c3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
c4	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
c5	0.0	0.95	0.0	0.0	1.0	0.0	0.6	0.0
c6	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.4
c7	0.0	0.65	0.0	0.0	0.6	0.0	1.0	0.0
c8	0.0	0.0	0.0	0.0	0.0	0.4	0.0	1.0

Figura 9. Relação de Similaridade para *displ*.

hp	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
c2	0.0	1.0	0.55	0.150...	0.75	0.0	0.099...	0.0
c3	0.0	0.55	1.0	0.6	0.300...	0.35	0.0	0.35
c4	0.0	0.150...	0.6	1.0	0.0	0.75	0.0	0.75
c5	0.0	0.75	0.300...	0.0	1.0	0.0	0.35	0.0
c6	0.0	0.0	0.35	0.75	0.0	1.0	0.0	1.0
c7	0.0	0.099...	0.0	0.0	0.35	0.0	1.0	0.0
c8	0.0	0.0	0.35	0.75	0.0	1.0	0.0	1.0

Figura 10. Relação de Similaridade para *hp*.

weight	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.0	0.042...	0.0	0.2925	0.0	0.96
c2	0.0	1.0	0.5	0.0	0.7775	0.0	0.517...	0.0
c3	0.0	0.5	1.0	0.0	0.722...	0.0	0.017...	0.0
c4	0.042...	0.0	0.0	1.0	0.0	0.75	0.0	0.002...
c5	0.0	0.7775	0.722...	0.0	1.0	0.0	0.295...	0.0
c6	0.2925	0.0	0.0	0.75	0.0	1.0	0.0	0.252...
c7	0.0	0.517...	0.017...	0.0	0.295...	0.0	1.0	0.0
c8	0.96	0.0	0.0	0.002...	0.0	0.252...	0.0	1.0

Figura 11. Relação de Similaridade para *weight*.

accel	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.25	0.0	0.0	0.25	0.0	0.0
c2	0.0	1.0	0.5	0.0	0.0	0.5	0.849...	0.050...
c3	0.25	0.5	1.0	0.0	0.0	1.0	0.650...	0.0
c4	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.949...
c5	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.949...
c6	0.25	0.5	1.0	0.0	0.0	1.0	0.650...	0.0
c7	0.0	0.849...	0.650...	0.0	0.0	0.650...	1.0	0.0
c8	0.0	0.050...	0.0	0.949...	0.949...	0.0	0.0	1.0

Figura 12. Relação de Similaridade para *accel*.

year	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.75	0.5	0.25	0.0	0.0	0.0	0.0
c2	0.75	1.0	0.75	0.5	0.25	0.0	0.0	0.0
c3	0.5	0.75	1.0	0.75	0.5	0.25	0.0	0.0
c4	0.25	0.5	0.75	1.0	0.75	0.5	0.25	0.0
c5	0.0	0.25	0.5	0.75	1.0	0.75	0.5	0.25
c6	0.0	0.0	0.25	0.5	0.75	1.0	0.75	0.5
c7	0.0	0.0	0.0	0.25	0.5	0.75	1.0	0.75
c8	0.0	0.0	0.0	0.0	0.25	0.5	0.75	1.0

Figura 13. Relação de Similaridade para *year*.

orig	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
c2	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
c3	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
c4	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
c5	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0
c6	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
c7	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0
c8	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0

Figura 14. Relação de Similaridade para *orig*.

mpg	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.8	1.0	0.0	0.6	0.0	0.9
c2	0.0	1.0	0.0	0.0	0.6	0.0	0.6	0.0
c3	0.8	0.0	1.0	0.8	0.0	0.8	0.0	0.7
c4	1.0	0.0	0.8	1.0	0.0	0.6	0.0	0.9
c5	0.0	0.6	0.0	0.0	1.0	0.0	1.0	0.0
c6	0.6	0.0	0.8	0.6	0.0	1.0	0.0	0.5
c7	0.0	0.6	0.0	0.0	1.0	0.0	1.0	0.0
c8	0.9	0.0	0.7	0.9	0.0	0.5	0.0	1.0

Figura 15. Relação de Similaridade para *mpg*.

### 5.2.3. Sin

A Figura 16 apresenta a relação de similaridade  $S_{in}$ .

Sin	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.107...	0.107...	0.256...	0.0	0.363...	0.0	0.351...
c2	0.107...	1.0	0.578...	0.164...	0.5325	0.071...	0.445...	0.078...
c3	0.107...	0.578...	1.0	0.228...	0.324...	0.228...	0.202...	0.085...
c4	0.256...	0.164...	0.228...	1.0	0.321...	0.5	0.107...	0.528...
c5	0.0	0.5325	0.324...	0.321...	1.0	0.107...	0.535	0.242...
c6	0.363...	0.071...	0.228...	0.5	0.107...	1.0	0.200...	0.521...
c7	0.0	0.445...	0.202...	0.107...	0.535	0.200...	1.0	0.178...
c8	0.351...	0.078...	0.085...	0.528...	0.242...	0.521...	0.178...	1.0

Figura 16. Relação de similaridade  $S_{in}$ .

#### 5.2.4. Sout

A Figura 17 apresenta a relação de similaridade  $S_{out}$ .

mpa	c1	c2	c3	c4	c5	c6	c7	c8
c1	1.0	0.0	0.8	1.0	0.0	0.6	0.0	0.9
c2	0.0	1.0	0.0	0.0	0.6	0.0	0.6	0.0
c3	0.8	0.0	1.0	0.8	0.0	0.8	0.0	0.7
c4	1.0	0.0	0.8	1.0	0.0	0.6	0.0	0.9
c5	0.0	0.6	0.0	0.0	1.0	0.0	1.0	0.0
c6	0.6	0.0	0.8	0.6	0.0	1.0	0.0	0.5
c7	0.0	0.6	0.0	0.0	1.0	0.0	1.0	0.0
c8	0.9	0.0	0.7	0.9	0.0	0.5	0.0	1.0

Figura 17. Relação de similaridade  $S_{out}$ .

#### 5.2.5. Sres

A Figura 18 apresenta a relação de similaridade  $S_{res}$ .

Sres	c1	c2	c3	c4	c5	c6	c7	c8
c1	1	0	1	1	0	1	0	1
c2	0	1	0	0	1	0	1	0
c3	1	0	1	1	0	1	0	1
c4	1	0	1	1	0	1	0	1
c5	0	1	0	0	1	0	1	0
c6	1	0	1	1	0	1	0	0
c7	0	1	0	0	1	0	1	0
c8	1	0	1	1	0	0	0	1

Figura 18. Relação de similaridade  $S_{res}$ .

#### 5.2.6. Clusters

A Figura 19 apresenta os resultados dos clusters máximos da base de casos são:  $H(B) = \{\{c_1, c_3, c_4, c_6\}, \{c_1, c_3, c_4, c_8\}, \{c_2, c_5, c_7\}\}$ .

Clusters							
1		3	4		6		
	2			5		7	
1		3	4				8

**Figura 19. Resultados dos procedimento *Clusters*.**

## 6. Conclusão

Este trabalho apresenta uma ferramenta de apoio ao método (CRR - Case Resemblance Relation), que se baseia numa relação de semelhança entre casos, para atender ao objetivo de solucionar problemas utilizando os casos ou situações já aprendidas.

## Referências

- [1] A. AAMODT and M. M. Richter. *Case-based reasoning foundations*. Cambridge University, 2006.
- [2] M. Abel. *Um Estudo Sobre Raciocínio Baseado em Casos*. Universidade Federal do Rio Grande do Sul, janeiro 1996.
- [3] H. M. Deitel and P. J. Deitel. *Java como Programar*. Pearson Prentice Hall, segunda edição edition, 2005. 1110p.
- [4] T. O. Fanoiki, Isabel Drummond, and Sandra Sandri. *Case Based Reasoning Retrieval and Reuse Using Case Resemblance Hypergraphs*, 2010.
- [5] R. C. Lee and W. M. Tepfenhart. *UML e C++ - Guia Prático de Desenvolvimento Orientado a Objeto*. Makron Books, São Paulo, SP, 2002. 550p.
- [6] A. C. Zimmermman, L. S. Encinas, J. M. Barreto, and L. O. Marin. Uma proposta de reutilização de de raciocínio baseado em casos - rbc em reconhecimento de faces humanas tri-dimensionais. II Congresso Brasileiro de Computação - CBComp2002.