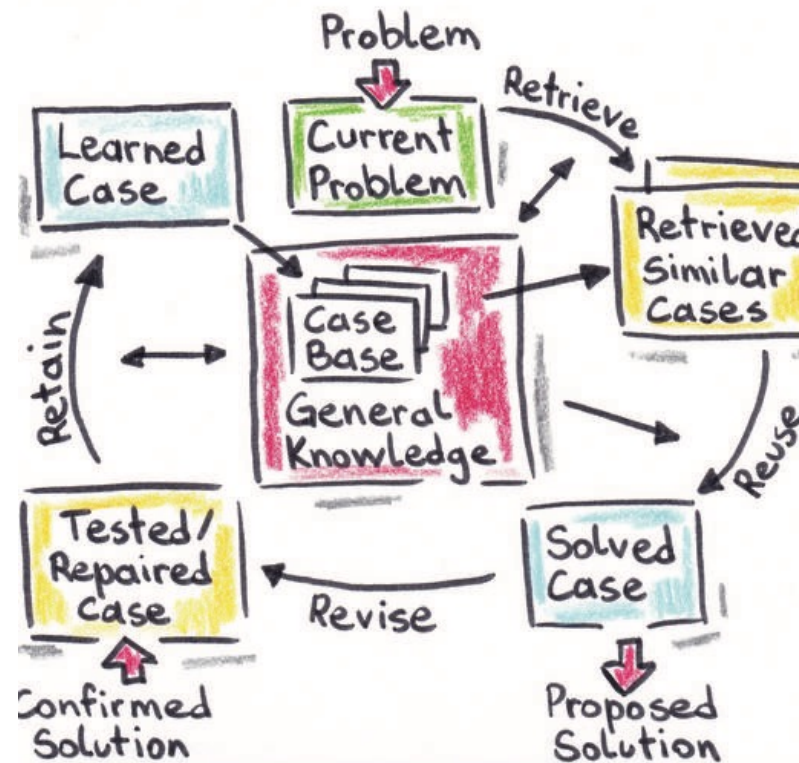


CONHECIMENTO E RACIOCÍNIO

3 – Raciocínio Baseado em Casos

(*CBR: Case based reasoning*)



Raciocínio Baseado em Casos

O Raciocínio Baseado em Casos (*Case Based Reasoning* ou CBR) é um paradigma (modelo) utilizado na implementação de muitos sistemas baseados em conhecimento:

- Baseia-se na representação de casos (ocorrências contextualizadas) que representam recordações (experiências passadas)
- Cada caso é composto por uma descrição e pela respetiva solução
- A resolução de um novo problema consiste em encontrar um caso passado semelhante ao atual e apresentar a sua solução diretamente ou adaptada ao novo contexto

Raciocínio Baseado em Casos

- Introduzido por Roger Schank nos anos 80

1º Sistema CBR

- CYRUS, Universidade de Yale, EUA, Janet Kolodner, 1983
Computerized Yale Retrieval and Updating System

Raciocínio Baseado em Casos

Protótipos acadêmicos

- **CHEF (Hammond, 1986):** Receitas culinárias novas a partir de outras conhecidas
- **PROTOS (Porter & Bareiss, 1986):** Diagnóstico de doenças auditivas
- **JUDGE (Bain, 1986):** Sentenças criminais (assassínio, assaltos, etc.)
- **PERSUADER (Sycara, 1987):** Resolução de conflitos laborais
- **HYPO (Ashley, 1988):** Direitos de autor e patentes: gera argumentos para defesa ou acusação baseado em casos anteriores
- **CASEY (Koton, 1989):** Diagnóstico de doenças de coração
- **JULIA (Hinrichs, 1989):** Projeto de ementas
- **PATDEX (Richter, 1991):** Diagnóstico de falhas em equipamento complexo
- **ARCHIE (Pearce et al., 1992):** Arquitetura: acede a projetos anteriores (criados por outros arquitetos)

Raciocínio Baseado em Casos

Aplicações reais e software comercial

- **KAIDARA Advisor**

- **CFM / BOEING:** Manutenção de Turbinas CFM 56-3
- **ESTALEIROS DE ODENSE:** Manutenção de Robots de Soldadura
- **ARIANE:** Fiabilidade dos Lançamentos
- **NATIONAL SEMICONDUCTOR:** Controlo de Qualidade em Semicondutores
- Manutenção do Metropolitano de Nápoles
- Diagnóstico de falhas em equipamento eletrónico para helicópteros e aviões (nomeadamente Airbus A330 e A340) na Sextant;
- Diagnóstico de falhas na Freightliner, USA, fabricante de veículos pesados do grupo DaimlerChrysler

- **Empolis Orenge / CBR-Works**

- **SIEMENS AG:** Suporte on-line a Sistemas de Automação
- **DEUTSCHE TELEKOM AG:** Helpdesk para Assistência Telefónica
- Bayer, Bosch, Lufthansa, Expo 2000 Hannover, Fujitsu-Siemens, Hewllet-Packard, Kluwer Academic Publishers, Volkswagen

Raciocínio Baseado em Casos

Princípio Básico do CBR :

usar a experiência passada para resolver um novo problema

- experiência passada: memória >> base de casos
- novo problema: novo caso

Raciocínio Baseado em Casos

Ao atender um novo paciente e analisar seus sintomas, o médico usa a sua experiência com outros pacientes *«Os problemas apresentados nos ouvidos do paciente são parecidos com um caso típico de otite média. Assim vou administrar-lhe um tratamento para otite média.»*.

Um técnico de reparação lembra-se de um defeito similar no tipo de máquina que está tentando reparar: *«Essa TV tem os mesmos problemas de uma que eu consertei na semana passada, então, também vou trocar as válvulas de saída de áudio.»*

Um arquiteto estuda as plantas de um prédio já existente ao planejar uma construção semelhante: *«No ano passado fiz uma casa de praia com três quartos, vou usar o plano daquele caso como uma base.»*

Um gestor de recursos humanos analisa candidaturas. Comparar perfis de candidatos com casos anteriores de funcionários bem-sucedidos na empresa para tomar decisões de contratação. *«Candidatos com perfil proativo e envolvimento em atividades extra curriculares têm normalmente mais sucesso na progressão de carreira»*

Raciocínio Baseado em Casos

- Todas estas situações têm em comum o facto de que, uma solução para um problema obtida no passado, foi reutilizada para guiar a solução do problema na situação presente.
- **Raciocínio Baseado em Casos** ou **Case Based Reasoning (CBR)** é a tecnologia de Inteligência Artificial inspirada neste modelo de cognição e comportamento humanos.

Raciocínio Baseado em Casos

- Armazenar experiência (casos) na memória
- Resolver novos problemas:
 - encontrar casos armazenados semelhantes ao novo problema
 - reutilizar (adaptar) os casos encontrados à nova situação
 - reutilização parcial ou completa
 - adaptar, ajustando diferenças pontuais
 - armazenar novo caso na memória

Raciocínio Baseado em Casos

Vantagens dos sistemas baseados no paradigma CBR:

- Funcionam de forma mais consentânea com o raciocínio pericial, i.e., baseiam-se em experiência passada
- A aprendizagem é simples: baseia-se na adição de "casos relevantes" à Base de Conhecimento (*Case Library*)
- A sua implementação evita a fase de extração de conhecimento (pode ser essencialmente baseada em documentos que descrevem casos passados)

Desvantagens:

- Dificuldade na adaptação de casos

Especialmente indicados para:

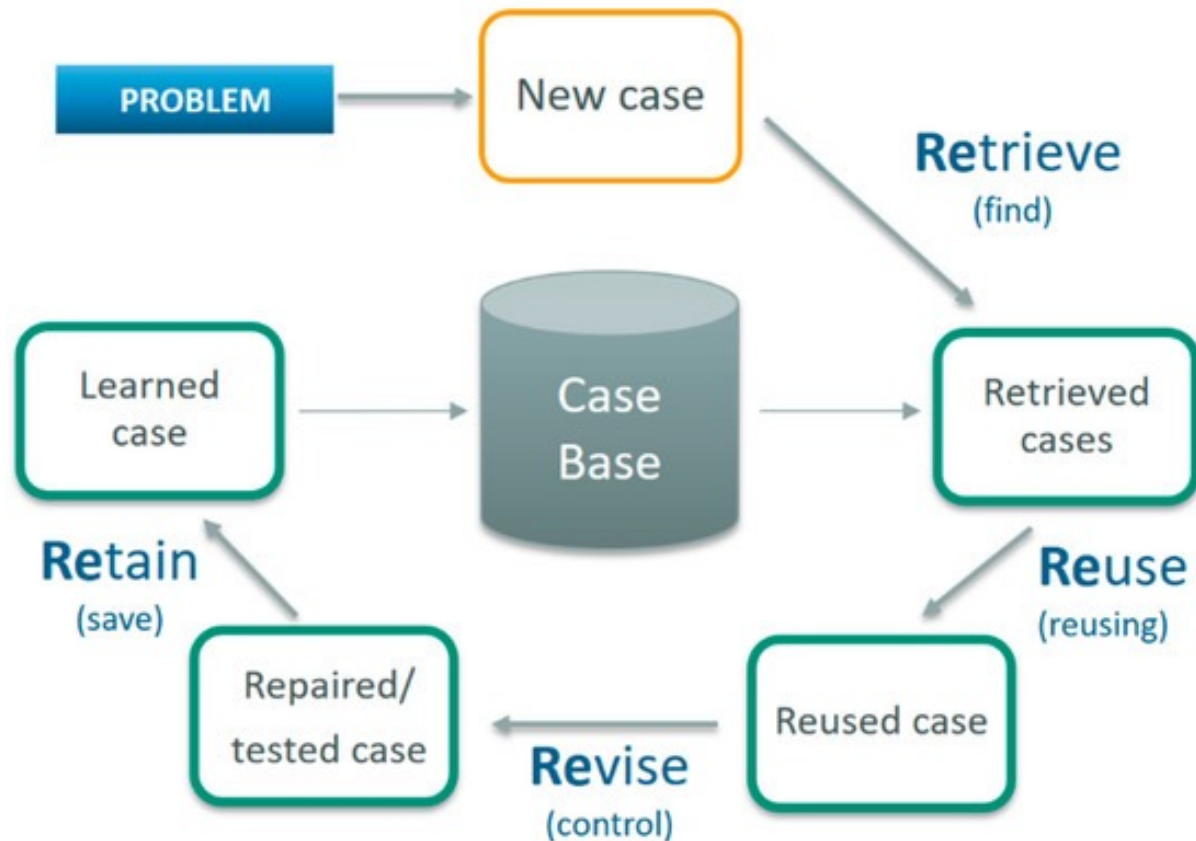
- Domínios que mudam com facilidade, dinâmicos
- Necessidade de aprendizagem constante e automática
- Domínios complexos e/ou mal compreendidos
- Domínios que dependam claramente de experiências passadas

Raciocínio Baseado em Casos

	Sistemas Baseados em Regras	Raciocínio Baseado em Casos
Domínio	Bem delimitado, bem compreendido, com alguma base, estável ao longo do tempo	Extenso, vago, mal compreendido, base teórica pobre, dinâmico ao longo do tempo
Representação do Conhecimento	Factos e regras de produção IF...THEN	Casos
Sistema proporciona	Respostas inferidas do zero	Respostas baseadas em acontecimentos precedentes
Explicação por	Traço das regras disparadas	Casos passados seleccionados
O sistema pode aprender	Geralmente não, regras modificadas manualmente, quando muito indução de regras a partir de exemplos (CBR ou regras ?)	Sim, por simples aquisição de novos casos (o sistema aprende com o seu uso)

Ciclo de um sistema CBR

- O **Ciclo de Aamodt & Plaza**, também conhecido como o **Ciclo de Engenharia do Conhecimento**, é um modelo que descreve as fases envolvidas no processo de resolução de problemas em sistemas de raciocínio baseado em casos (CBR). Este ciclo foi proposto pelos investigadores Agnar Aamodt e Enric Plaza.



Ciclo de um sistema CBR

1. **Recuperação (Retrieve):** Nesta fase, o sistema identifica casos relevantes que estão armazenados na base de casos. A recuperação envolve a comparação do problema atual com casos passados para encontrar aqueles que são mais similares.
2. **Reutilização (Reuse):** Após a recuperação dos casos, o sistema reutiliza a solução ou parte da solução dos casos recuperados. Isso implica em adaptar ou modificar a solução para se adequar ao problema atual, se necessário.
3. **Revisão (Revise):** Nesta fase, a solução proposta é examinada e avaliada. Pode haver a necessidade de ajustes ou modificações para garantir que a solução seja apropriada para o problema atual.
4. **Retenção (Retain):** Finalmente, a solução revisada é incorporada à base de casos. Isso aumenta o conhecimento do sistema, tornando-o mais capaz de lidar com problemas futuros semelhantes (**aprendizagem**).

Raciocínio Baseado em Casos

EXEMPLO

- Exemplo: diagnóstico de falhas em automóveis
 - **Diagnóstico da falha**
 - Perito observa os **sintomas** da avaria
 - Motor não arranca
 - Perito mede valores
 - Voltagem da bateria = 6.3 V
 - **Objetivo:**
 - Determinar a causa da avaria
 - Bateria descarregada
 - Propor uma solução:
 - Trocar bateria

Raciocínio Baseado em Casos

EXEMPLO

- Exemplo: diagnóstico de falhas em automóveis
 - Um **caso** é composto por:
 - Sintomas da avaria
 - Avaria detetada
 - Causa da avaria
 - Reparação adequada
 - Os casos são guardados na base de casos (case library)
 - **Objetivo:** quando surge uma nova avaria, encontrar na base de casos as situações similares armazenadas

Raciocínio Baseado em Casos

EXEMPLO: Base de Casos (Case Base, Case Library)

- Caso:
 - representa tipicamente a descrição de uma situação (problema) conjuntamente com as experiências adquiridas (solução) durante a sua resolução

Caso 1	ATRIBUTO	VALOR
Sintomas:	Problema:	faróis dianteiros não funcionam
	Modelo:	Chevy Sierra 1500
	Ano:	2005
	Voltagem da bateria:	13.6 V
	Estado dos faróis:	não danificados
	Estado do interruptor dos faróis:	a funcionar
Solução:	Diagnóstico: avaria nos fusíveis dos faróis	
	Ação: trocar fusíveis dos faróis	

Raciocínio Baseado em Casos

EXEMPLO: Base de Casos (Case Base, Case Library)

- Outro Caso:
 - Casos são independentes
 - Não permitem generalização

Caso 2	ATRIBUTO	VALOR
Sintomas:	Problema:	faróis dianteiros não funcionam
	Modelo:	Tesla S
	Ano:	2018
	Voltagem da bateria:	12.9 V
	Estado dos faróis:	ótica partida
	Estado do interruptor dos faróis:	a funcionar
Solução:	Diagnóstico: faróis dianteiros partidos	
	Ação: trocar faróis dianteiros	

Base de Casos (Case Base, Case Library)

- A *Case Library* consiste num **conjunto de casos** que armazenem diferentes situações para um determinado problema:

Caso 1	ATRIBUTO	VALOR
Sintomas:	Problema:	faróis dianteiros não funcionam
	Modelo:	<u>Chevy Sierra 1500</u>
	Ano:	2005
	Voltagem da bateria:	13.6 V
	Estado dos faróis:	não danificados
	Estado do interruptor dos faróis:	a funcionar
Solução:	Diagnóstico: avaria nos fusíveis dos faróis	
	Ação: trocar fusíveis dos faróis	

Caso 2	ATRIBUTO	VALOR
Sintomas:	Problema:	faróis dianteiros não funcionam
	Modelo:	Tesla S
	Ano:	2018
	Voltagem da bateria:	12.9 V
	Estado dos faróis:	ótica partida
	Estado do interruptor dos faróis:	a funcionar
Solução:	Diagnóstico: faróis dianteiros partidos	
	Ação: trocar faróis dianteiros	

Raciocínio Baseado em Casos

EXEMPLO

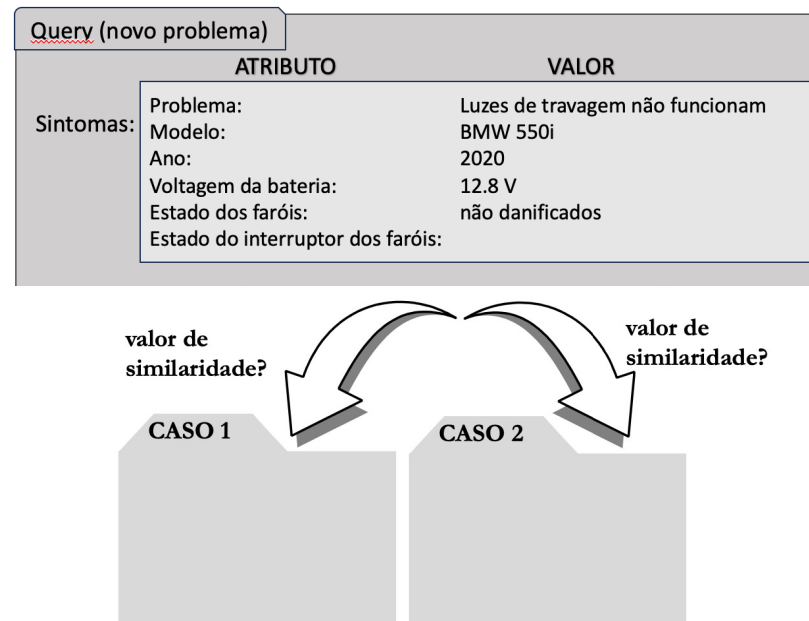
- Novo problema: contém apenas sintomas.
- Objetivo: encontrar uma solução
- Pode haver sintomas sem valores definidos

Query (novo problema)		
ATRIBUTO		VALOR
Sintomas:	Problema:	Luzes de travagem não funcionam
	Modelo:	BMW 550i
	Ano:	2020
	Voltagem da bateria:	12.8 V
	Estado dos faróis:	não danificados
	Estado do interruptor dos faróis:	

CBR: fase de Retrieve

Similaridade global

- Escolher o caso da *Case Library* mais semelhante ao problema atual
- Adaptar o caso escolhido para a resolver o novo problema
- É necessário calcular uma **métrica de similaridade** entre os casos conhecidos e o novo caso
 - SIMILARIDADE GLOBAL:
 - calculada a partir de valores de semelhança entre os vários atributos.
 - Cada atributo pode ter um **peso** diferente (importância)



CBR: fase de Retrieve

Similaridade global

- Similaridades entre atributos (similaridade local)
 - Faróis dianteiros não funcionam vs. Luzes de travagem não funcionam **0.8**
 - Faróis dianteiros não funcionam vs. Motor não arranca **0.4**
 - Voltagem da bateria:
 - 12.6 V vs. 13.6 V **0.9**
 - 12.6 V vs. 6.7 V **0.1**

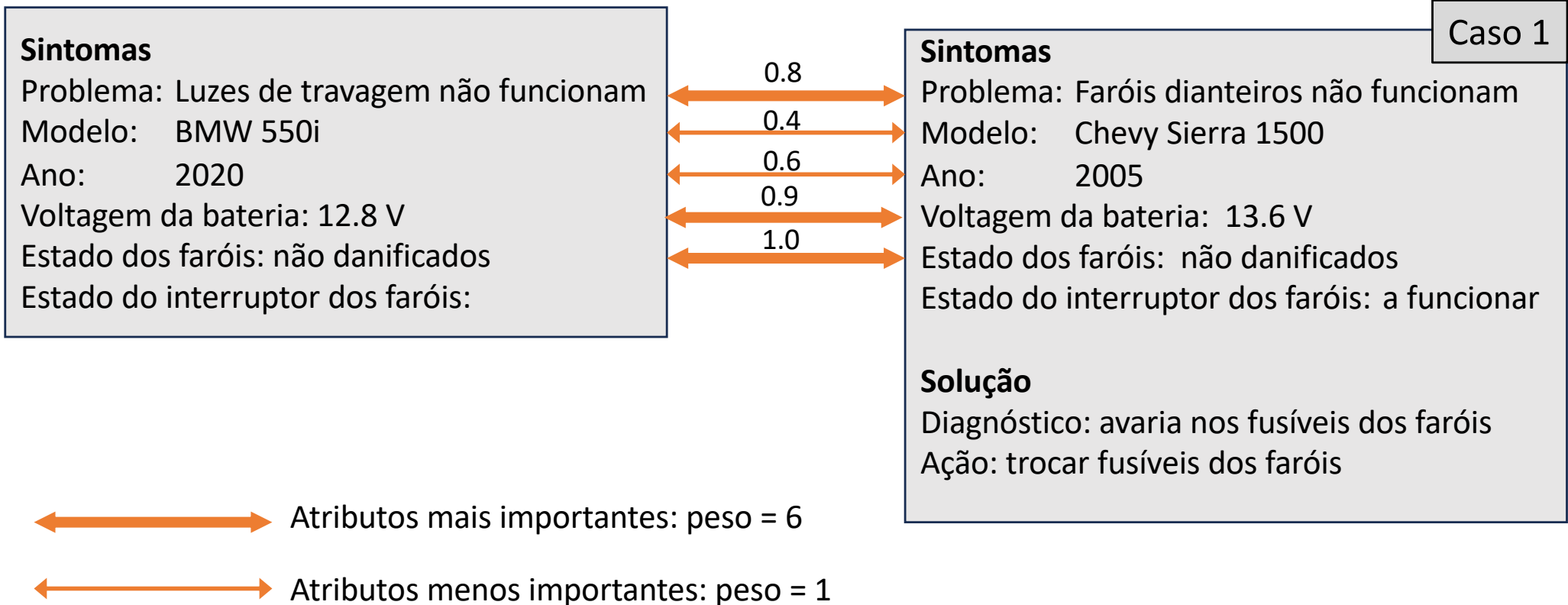
CBR: fase de Retrieve

Similaridade global

- Atributos em têm diferentes relevâncias
 - Mais importantes
 - Problema
 - Voltagem da bateria
 - Estado dos faróis
 - Estado do interruptor dos faróis
 - Menos importantes
 - Modelo
 - Ano
- Os atributos devem ter pesos associados

CBR: fase de Retrieve

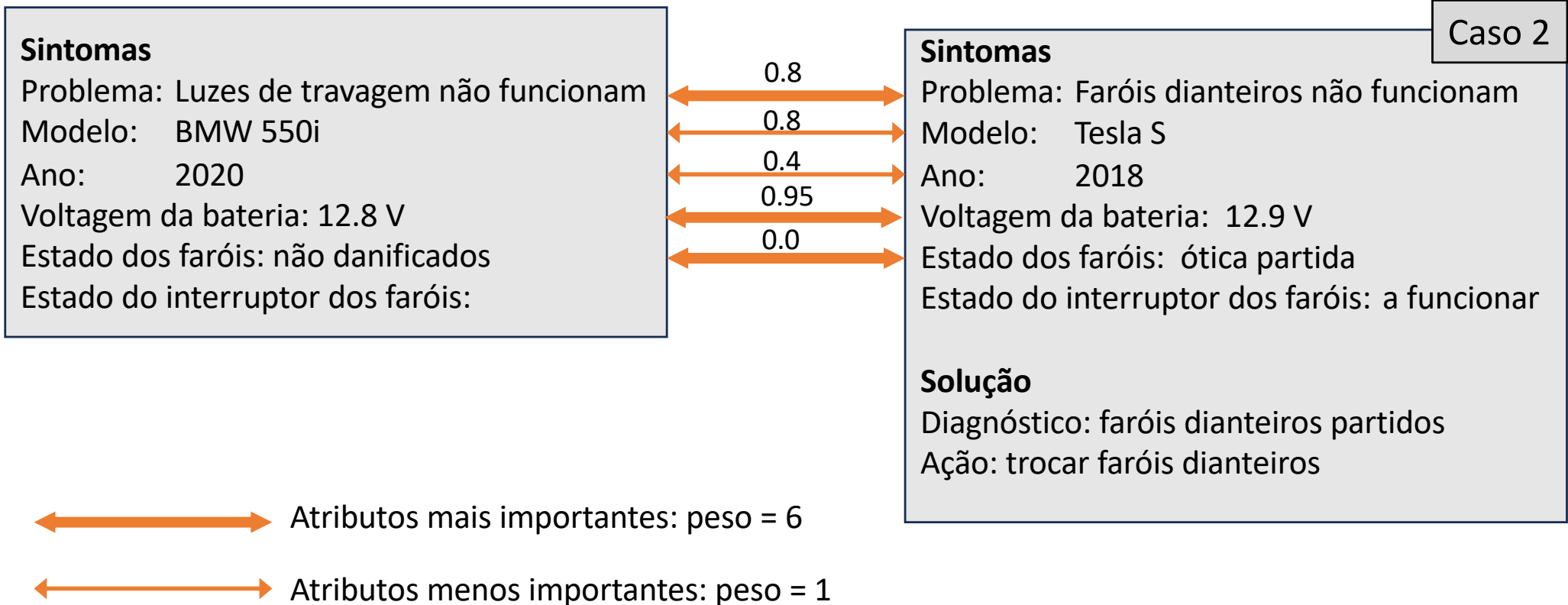
Similaridade global



$$\text{Similaridade} = [(6 * 0.8) + (1 * 0.4) + (1 * 0.6) + (6 * 0.9) + (6 * 1.0)] / 20 = 0.86$$

CBR: fase de Retrieve

Similaridade global



$$\text{Similaridade} = [(6 * 0.8) + (1 * 0.8) + (1 * 0.4) + (6 * 0.95) + (6 * 0.0)] / 20 = 0.585$$

CBR: fase de Reuse

- Propor a solução de um caso passado como solução para o caso atual (*null adaptation*)
OU
- Adaptar a solução de um caso passado ao caso atual:
 - **Adaptação Transformacional:** são definidos operadores transformacional: por exemplo regras, algoritmos, expressões matemáticas - que guiam a adaptação da solução passada ao caso atual
 - **Adaptação Derivacional:** um método usado no passado para adaptar uma solução é reutilizado; a cada caso tem de ser associada alguma forma de descrição desses métodos

CBR: fase de Reuse

Adaptação do novo caso usando o caso escolhido

O Caso 1 tem similaridade mais alta
É criado um novo caso adaptando o novo problema

Novo caso: Caso 3

Sintomas

Problema: Luzes de travagem não funcionam

Modelo: BMW 550i

Ano: 2020

Voltagem da bateria: 12.8 V

Estado dos faróis: não danificados

Estado do interruptor dos faróis:

Solução

Diagnóstico: avaria nos fusíveis das luzes de travagem

Ação: trocar fusíveis dos luzes de travagem

Caso 1

Sintomas

Problema: Faróis dianteiros não funcionam

Modelo: Chevy Sierra 1500

Ano: 2005

Voltagem da bateria: 13.6 V

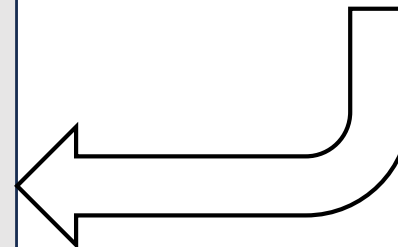
Estado dos faróis: não danificados

Estado do interruptor dos faróis: a funcionar

Solução

Diagnóstico: avaria nos fusíveis dos faróis

Ação: trocar fusíveis dos faróis



CBR: fase de Revise

- **Avaliação** da solução proposta
 - Por um **perito externo**:
 - Exemplo: um técnico ou médico confirma ou não a solução proposta pelo sistema
 - Por um **processo de simulação**:
 - CHEF - As receitas culinárias propostas pelo SP são aplicadas a um simulador que tenta validá-las
- Eventual **reparação** dessa solução
 - Exemplo: o diagnóstico proposto é modificado pelo utilizador que pode também comunicar um novo ou escolher um já existente mas não associado ao contexto atual

CBR: fase de Retain

- Que conhecimento reter ?
- Quando o reter ?
- Como o reter?
- O caso atual deverá ser inserido como **novo caso** ?
- Deverão ser feitas apenas algumas "**atualizações**" na base de casos ?
 - A decisão é tomada em função do conteúdo do caso atual e do conteúdo dos casos que já figuram na biblioteca;
 - A aprendizagem realiza-se quer por erro, quer por sucesso, pelo que também é importante considerar o desfecho obtido no processamento

CBR: medidas de similaridade

- Usadas na Fase de **Retrieve**
- Similaridades Locais
 - Medem a semelhança entre pares de atributos
- Similaridade Global
 - Mede a semelhança de cada caso da Case Library com o novo caso
 - Métrica pesada: cada atributo tem um peso associado

CBR: fase de Retrieve

Similaridade global

- Valor de similaridade é expresso como número real entre 0.0 (nenhuma semelhança) e 1.0 (igualdade)
- Valor de similaridade é calculado para cada caso na base de casos conforme os valores dos atributos.
- É uma métrica pesada: cada atributo pode ter um peso diferente
- Pode ser calculada usando a **distancia linear** ou a **distância euclidiana** entre atributos
- O valor da semelhança global, S_{Global} , é normalmente definido como a diferença de uma **distância normalizada**, D , para 1 :
$$S_{Global} = 1 - D$$
- **RETRIEVE** : devolve os casos com similaridade global acima de um limiar estabelecido

CBR: fase de Retrieve

Similaridade global

Por exemplo, numa *case library* com 5 casos, após o cálculo da similaridade global:

Caso	Distância Linear pesada	Similaridade global
1	0.90	0.10
2	0.10	0.90
3	0.40	0.60
4	0.80	0.20
5	0.05	0.95

RETRIEVE: LIMIAR (*Threshold*) ≥ 0.90

Caso retrieved	Similaridade global
2	0.90
5	0.95

CBR: fase de Retrieve

Similaridade global

Distância Linear para n atributos

$$D_L(q, s) = \frac{\sum_{i=1}^n d(q_i, s_i) \times w_i}{\sum_{i=1}^n w_i}$$

Distância Euclidiana para n atributos

$$D_E(q, s) = \sqrt{\frac{\sum_{i=1}^n [d(q_i, s_i)]^2 \times w_i}{\sum_{i=1}^n w_i}}$$

D_L Distância linear global entre casos

D_E Distância euclidiana global entre casos

q Caso novo

s Caso da biblioteca

q_i Valor do atributo de índice i no caso novo

s_i Valor do mesmo atributo no caso da biblioteca

w_i Fator de relevância (peso) para o atributo de índice i

d Função **distância local** entre atributos correspondentes

CBR: fase de Retrieve

distância local

- Atributos podem ter tipos de dados diferentes:
 - Valores numéricos
 - Valores simbólicos (string, expressões)
 - Datas
 - Coordenadas geográficas
 - Valores booleanos
 - ...
- A distância local é calculada para cada par de atributos:
casos conhecidos / novo caso
- Muitas vezes é necessário um **perito** no conhecimento para estabelecer as similaridades entre pares de símbolos
- As distâncias locais devem ser **normalizadas**

CBR: fase de Retrieve

Distância local: atributos numéricos

- Distância linear
- Distância euclidiana
- Normalização
- Importância da simetria das distâncias

CBR: fase de Retrieve

Distância local: atributos numéricos

- Simetria: permite dar mais importância aos valores “por baixo” ou “por cima”
- Normalização de atributos numéricos:

$$V = V / \max(V_{i(i=1 \dots \text{Num casos})})$$

CBR: fase de Retrieve

Distância local: atributos numéricos

Distância linear entre atributos numéricos

Atributo: preço;

Valor: número real

Preço caso conhecido 1 (c_1): 900

Preço caso conhecido 2 (c_2): 1100

Preço novo caso (q): 1000

- $D_L(c, q) = |v_c - v_q|$

$$D_L(c_1, q) = |900 - 1000| = 100$$

$$D_L(c_2, q) = |1100 - 1000| = 100$$

- D_L normalizada

$$D_L(c_1, q) = |0.82 - 0.91| = 0.09$$

$$D_L(c_2, q) = |1 - 0.91| = 0.09$$

CBR: fase de Retrieve

Distância local: atributos numéricos

- **Distância euclidiana** entre atributos numéricos

- $D_E(c, q) = \sqrt{(c - q)^2}$

$$D_E(c_1, q) = \sqrt{(900 - 1000)^2} = 100$$

$$D_E(c_2, q) = \sqrt{(1100 - 1000)^2} = 100$$

- D_E normalizada

$$D_E(c_1, q) = \sqrt{(0.82 - 0.91)^2} = 0.09$$

$$D_E(c_2, q) = \sqrt{(1 - 0.91)^2} = 0.09$$

CBR: fase de Retrieve

Distância local: atributos numéricos

- Exemplo: Preço novo caso (q): 1000
 - Normalizando: $1000/1600 = 0,625$

- Case Library:

Caso	Atributo preço	valor normalizado	Distâncias lineares normalizadas
1	700	$700/1600 = 0,4375$	$ 0,625 - 0,4375 = 0.1875$
2	1000	$1000/1600 = 0,625$	$ 0,625 - 0,625 = 0$
3	500	$500/1600 = 0,3125$	$ 0,625 - 0,3125 = 0.3125$
4	1600	$1600/1600 = 1$	$ 0,625 - 1 = 0.375$
5	1200	$1200/1600 = 0,75$	$ 0,625 - 0,75 = 0.125$

CBR: fase de Retrieve

Distância local: atributos simbólicos (strings)

- Alguns atributos podem ter valores atribuídos como strings dentro de uma lista limitada (enumerações)
- As distâncias locais são definidas a partir de tabelas de similaridades
- As tabelas definem as similaridades entre pares de atributos
- Os valores são atribuídos manualmente pelo perito no domínio de conhecimento

CBR: fase de Retrieve

Distância local: atributos simbólicos (strings)

- Exemplo: atributo “carroçaria” numa base de dados de um stand de venda de carros:
 - Lista enumerada do tipo de carroçaria:
 - *Hatchback, Berlina, Carrinha, Descapotável, Coupé*
- Tabela de similaridade (TS):

	<i>Hatchback</i>	<i>Berlina</i>	<i>Carrinha</i>	<i>Descapotável</i>	<i>Coupé</i>
<i>Hatchback</i>	1.0	0.8	0.7	0.1	0.0
<i>Berlina</i>	0.8	1.0	0.8	0.2	0.1
<i>Carrinha</i>	0.7	0.8	1.0	0.1	0
<i>Descapotável</i>	0.1	0.2	0.1	1.0	0.8
<i>Coupé</i>	0.0	0.1	0.0	0.8	1.0

CBR: fase de Retrieve

Distância local: atributos enumerados (strings)

- Cálculo da distância local para atributos deste tipo:

- **Caso novo (q): atributo carroçaria, valor *Berlina***
- Caso 1 da CL (c1): atributo carroçaria, valor *Carrinha*
- Caso 2 da CL (c2): atributo carroçaria, valor *Berlina*
- Caso 3 da CL (c3): atributo carroçaria, valor *Coupé*
- Caso 4 da CL (c4): atributo carroçaria, valor *Descapotável*
- Caso 5 da CL (c5): atributo carroçaria, valor *Carrinha*

$$D(c_1, q) = 1 - TS(Carrinha, Berlina) = 1 - 0.8 = 0.2$$

$$D(c_2, q) = 1 - TS(Berlina, Berlina) = 1 - 1 = 0.0$$

$$D(c_3, q) = 1 - TS(Coupé, Berlina) = 1 - 0.1 = 0.9$$

$$D(c_4, q) = 1 - TS(Descapotável, Berlina) = 1 - 0.1 = 0.9$$

$$D(c_5, q) = 1 - TS(Carrinha, Berlina) = 1 - 0.8 = 0.2$$

CBR: fase de Retrieve

Distância local: atributos com datas

- Cálculo da distância local para atributos deste tipo:
 - Por exemplo, considerar apenas o ano >> passa a atributo numérico
 - **Caso novo (q): O cliente pretende um carro de 2022**
 - Caso 1 da CL (c1): atributo *data_registro*, valor *31-03-2016*
 - Caso 2 da CL (c2): atributo *data_registro*, valor *01-01-2024*
 - Caso 3 da CL (c3): atributo *data_registro*, valor *20-09-2022*
 - Caso 4 da CL (c3): atributo *data_registro*, valor *01-12-2010*
 - Caso 5 da CL (c3): atributo *data_registro*, valor *31-03-2002*

CBR: fase de Retrieve

Distância local: atributos com datas

- Cálculo da distância local considerando apenas o ano
 - Q = 2022 (normalizado: 0.999)

Caso	Ano	valor normalizado	Distâncias lineares normalizadas
1	2016	0.996	0.003
2	2024	1	0.001
3	2022	0.999	0
4	2010	0.993	0.006
5	2002	0.9891	0.00987

CBR: fase de Retrieve

Um exemplo

Case Library: venda de carros usados

caso	preço	carroçaria	ano
1	700	Carrinha	2016
2	1000	Berlina	2024
3	500	Coupé	2022
4	1600	Descapotável	2010
5	1200	Carrinha	2002

Cliente pretende (novo caso q): *Berlina, 1000 eur, ano 2022*

CBR: fase de Retrieve

Um exemplo

Na fase de **Retrieve** calculam-se as similaridades globais de todos os casos

Vamos assumir os pesos w_i : preço = 5; carroçaria = 2; ano = 1

$$D_L(q, s) = \frac{\sum_{i=1}^n d(q_i, s_i) \times w_i}{\sum_{i=1}^n w_i}$$

caso	D_L preço	D_L carroçaria	D_L ano
1	0.1875	0.2	0.003
2	0	0	0.001
3	0.3125	0.9	0
4	0.375	0.9	0.006
5	0.125	0.2	0.00987

CBR: fase de Retrieve

Um exemplo

Na fase de **Retrieve** calculam-se as similaridades globais de todos os casos

Vamos assumir os pesos w_i : **preço = 5; carroçaria = 2; ano = 1**

$$D_L(q, s) = \frac{\sum_{i=1}^n d(q_i, s_i) \times w_i}{\sum_{i=1}^n w_i}$$

caso	D_L	Similaridade = $1 - D_L$
1	0.1676	0.8324
2	0.0001	0.9999
3	0.4203	0.5797
4	0.4601	0.5399
5	0.1294	0.8706

CBR: fase de Retrieve

Um exemplo

Na fase de **REUSE** analisa-se o caso indicado pela fase de RETRIEVE.

É necessária alguma adaptação?

2	1000	Berlina	2024	Peugeot 508
---	------	---------	------	-------------

caso	preço	carroçaria	ano	Carro disponível
1	700	Carrinha	2016	BMW 520D
2	1000	Berlina	2024	Peugeot 508
3	500	Coupé	2022	Fiat Coupe
4	1600	Descapotável	2010	BMW Z3
5	1200	Carrinha	2002	Toyota Auris

CBR: fase de Retrieve

Um exemplo

Na fase de **REVISE** analisa-se o caso indicado pela fase de RETRIEVE e adaptado na fase de REUSE.

É necessária alguma alteração, revisão dos atributos?

2	1000	Berlina	2024	Peugeot 508
---	------	---------	------	-------------

caso	preço	carroçaria	ano	Carro disponível
1	700	Carrinha	2016	BMW 520D
2	1000	Berlina	2024	Peugeot 508
3	500	Coupé	2022	Fiat Coupe
4	1600	Descapotável	2010	BMW Z3
5	1200	Carrinha	2002	Toyota Auris

CBR: fase de Retrieve

Um exemplo

Na fase de **RETAIN** decide-se se o novo caso é acrescentado à case library.

Nesta situação não se justifica uma vez que não há adição de nova informação

2	1000	Berlina	2024	Peugeot 508
---	------	---------	------	-------------

caso	preço	carroçaria	ano	Carro disponível
1	700	Carrinha	2016	BMW 520D
2	1000	Berlina	2024	Peugeot 508
3	500	Coupé	2022	Fiat Coupe
4	1600	Descapotável	2010	BMW Z3
5	1200	Carrinha	2002	Toyota Auris

Raciocínio Baseado em Casos

EXERCÍCIOS

CBR: Exercício 1

Assuma que possui uma Case Library com perfis de compra de telemóveis

Assuma os pesos indicados a vermelho

caso	Idade (3)	Profissão (2)	Rendimento (5)	Modelo
1	12	Estudante	100	i-phone 11
2	14	Estudante	200	i-phone 11
3	18	Estagiário	800	i-phone 12
4	38	Médico	1500	i-phone 14
5	45	Advogado	2000	i-phone 15

Execute a fase de **Retrieve** calculando a similaridade linear global

Que modelo sugerir a um novo utilizador com este perfil?

16 anos, Estudante, rendimento 300

Que modelo sugerir a um novo utilizador com este perfil?

50 anos, Médico, rendimento 5000

CBR: Exercício 1

Assuma que possui uma Case Library com perfis de compra de telemóveis

Assuma os pesos indicados a vermelho

caso	Idade (3)	Profissão (2)	Rendimento (5)	Modelo
1	12	Estudante	100	i-phone 11
2	14	Estudante	200	i-phone 11
3	18	Estagiário	800	i-phone 12
4	38	Médico	1500	i-phone 14
5	45	Advogado	2000	i-phone 15

Query: 16 anos, Estudante, rendimento 300

CALCULAR:

DL para a idade

DL para a Profissão

DL para o rendimento

CBR: Exercício 1

Atributo idade: distância linear normalizada

$Q = 16 \gg 16/45 = 0,36$

caso	Idade	Normalizada	DL
1	12	$12/45 = 0,27$	0,089
2	14	$14/45 = 0,31$	0,044
3	18	$18/45 = 0,40$	0,044
4	38	$38/45 = 0,84$	0,489
5	45	$45/45 = 1,00$	0,644

Query: 16 anos, Estudante, rendimento 300


CBR: Exercício 1

Atributo profissão:

1º) Criar tabela de similaridades

2º) Calcular a distância local

Q = Estudante



	Estudante	Estagiário	Médico	Advogado
Estudante	1,0	0,8	0,0	0,0
Estagiário	0,8	1,0	0,0	0,0
Médico	0,0	0,0	1,0	0,9
Advogado	0,0	0,0	0,9	1,0

caso	Profissão	Distância local
1	Estudante	$1 - 1 = 0$
2	Estudante	$1 - 1 = 0$
3	Estagiário	$1 - 0,8 = 0,2$
4	Médico	$1 - 0 = 1$
5	Advogado	$1 - 0 = 1$

CBR: Exercício 1


Atributo rendimento: distância linear normalizada

$$Q = 300 \gg 300/2000 = 0,36$$

caso	Rendimento	Normalização	DL
1	100	$100/2000 = 0,05$	0,1
2	200	$200/2000 = 0,10$	0,05
3	800	$800/2000 = 0,40$	0,25
4	1500	$1500/2000 = 0,75$	0,60
5	2000	$2000/2000 = 1,00$	0,85

Query: 16 anos, Estudante, rendimento 300

CBR: Exercício 1

$$D_L(q, s) = \frac{\sum_{i=1}^n d(q_i, s_i) \times w_i}{\sum_{i=1}^n w_i}$$


Cálculo das similaridades GLOBAIS

Assuma os pesos indicados a vermelho

caso	DL Idade (3)	DL Profissão (2)	DL Rendimento (5)	DISTÂNCIA LINEAR GLOBAL PESADA	SIMILARIDADE GLOBAL
1	0,089	0	0,1	0,08	0,92
2	0,044	0	0,05	0,04	0,96
3	0,044	0,2	0,25	0,18	0,82
4	0,489	1	0,60	0,65	0,35
5	0,644	1	0,85	0,82	0,18

CBR: Exercício 1

Assuma que possui uma Case Library com perfis de compra de telemóveis

Assuma os pesos indicados a vermelho

caso	Idade (3)	Profissão (2)	Rendimento (5)	Modelo
1	12	Estudante	100	i-phone 11
2	14	Estudante	200	i-phone 11
3	18	Estagiário	800	i-phone 12
4	38	Médico	1500	i-phone 14
5	45	Advogado	2000	i-phone 15

Que modelo sugerir a um novo utilizador com este perfil?

16 anos, Estudante, rendimento 300

>>> Sugestão baseada em perfis semelhantes: Um i-phone 11