

Universidade do Minho

Escola de Engenharia

Departamento de Informática

Sistemas de Aprendizagem

Aprendizagem e Extração de Conhecimento
Perfil Sistemas Inteligentes @ MiEI/4º – 1º Semestre

Cesar Analide, José Neves, Paulo Novais

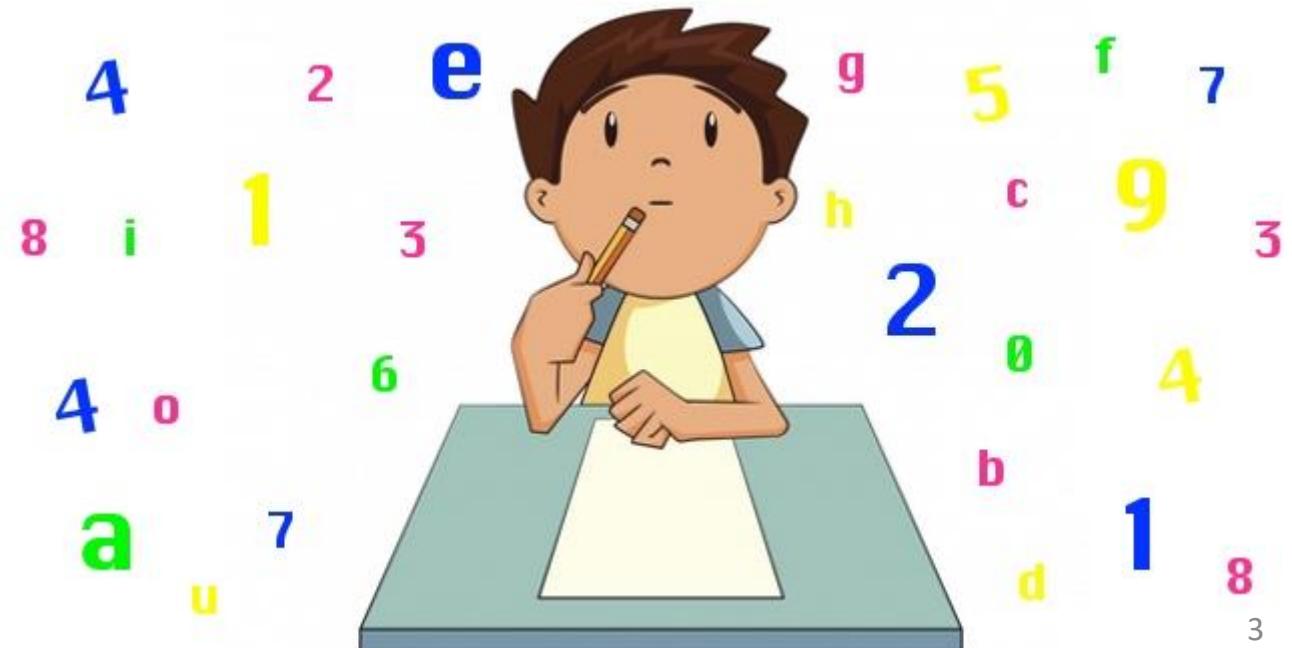
Definição de Aprendizagem

- Ato de entender e absorver conceitos;
- Adquirir conhecimento;
- ...



Definição de Aprendizagem

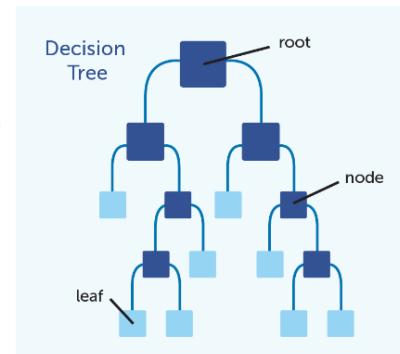
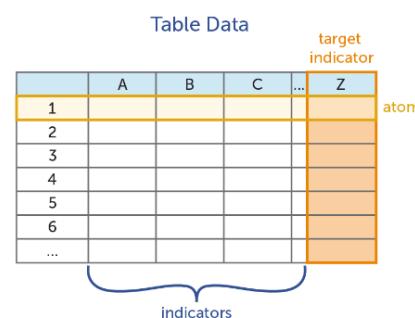
- Aprendizagem (<http://priberam.pt/dlpo/aprendizagem>):
 - Ato ou efeito de aprender; Tempo durante o qual se aprende; Experiência que tem quem aprendeu.





- Aprendizagem Simbólica;
- Redes Neuronais Artificiais;
- Raciocínio Baseado em Casos;
- Aprendizagem por Reforço;
- Árvores de Decisão;
- Algoritmos Genéticos e Evolucionários;
- Máquinas de Vetores de Suporte
- Inteligência de Grupo;
- e muitos outros...

Exemplos de Sistemas de Aprendizagem



Universidade do Minho

Escola de Engenharia

Departamento de Informática

Raciocínio Baseado em Casos

Case Based Reasoning

Aprendizagem e Extração de Conhecimento
Perfil Sistemas Inteligentes @ MiEI/4º – 1º Semestre

Cesar Analide, Filipe Gonçalves

- Raciocínio Baseado em Casos
- “[...]those who cannot remember the past,
are condemned to repeat it[...]
Carlos Santayana
- “[...]quem não for capaz de recordar o passado,
está condenado a repeti-lo[...]

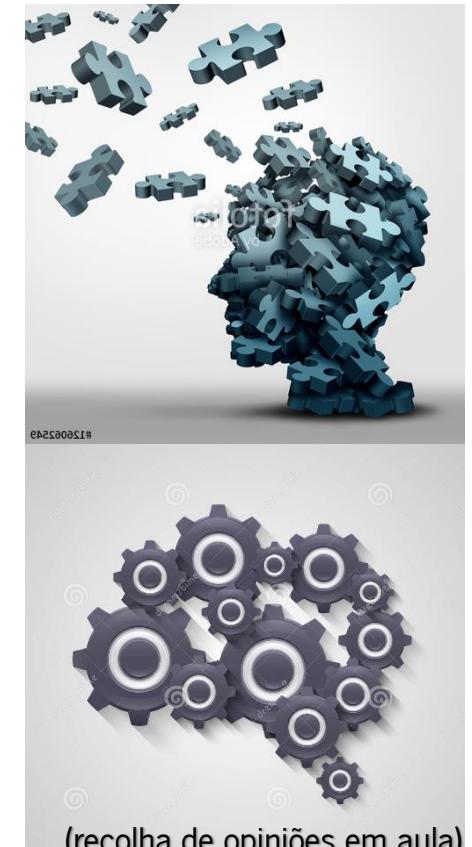


- O que é RBC e porquê raciocinar com base em experiências?
- Raciocinar com Casos
- Modelo
- Posicionamento em relação a outros métodos
- Aplicabilidade
- Futuras direções
- Assistentes pessoais digitais
- Bibliografia recomendada
- Exemplos práticos e exercícios



O que é Raciocínio Baseado em Casos (RBC)?

- Resolver problemas utilizando soluções prévias,



O que é Raciocínio Baseado em Casos (RBC)?

- Definição:

- RBC é raciocinar por recordação.

- Pressuposto:

- Problemas similares terão soluções similares.



© MARK ANDERSON



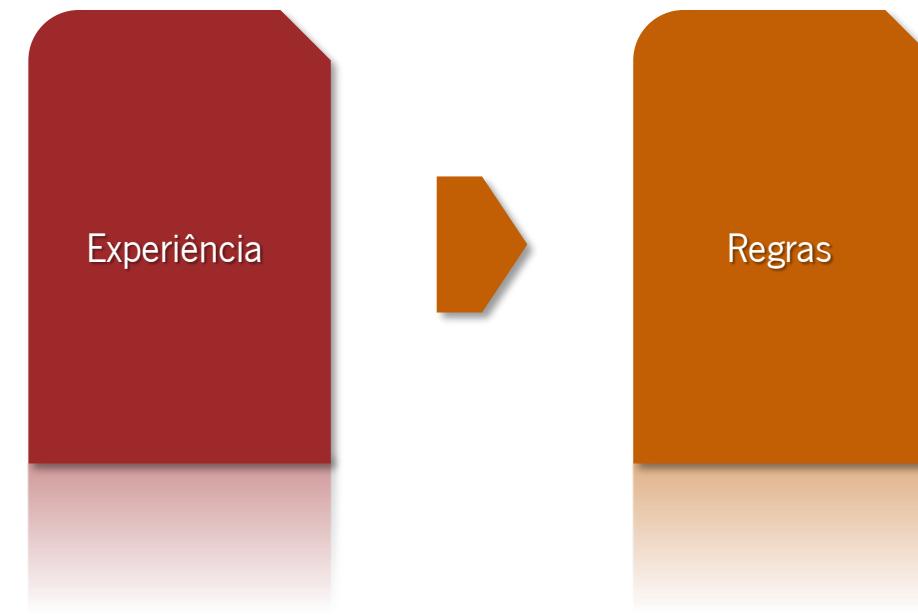
"Objection! Elephants never forget!"

Nós somos naturais sistemas RBC

- Um médico define uma terapêutica para o tratamento de uma doença **usando conhecimento de diagnósticos e tratamentos feitos no passado**;
- Um mecânico **experiente**, face à observação dos ruídos de um veículo, identifica o problema e determina a reparação a efetuar, **com base em soluções que adaptou no passado para problemas semelhantes**;
- Para a atribuição de crédito a clientes, uma entidade bancária **compara** cada caso com outros casos passados, que envolveram empresas ou indivíduos em **situações financeiras semelhantes, atribuindo o crédito de acordo com as decisões tomadas anteriormente**;
- Um vendedor imobiliário propõe **soluções idênticas para clientes de classes/faixas sociais próximas**;
- **Jurisprudência** é o conjunto de decisões sobre interpretações das leis feitas pelos tribunais.

Porquê RBC?

- Porque um dos problemas mais frequente na construção de Sistemas Inteligentes provém da dificuldade em **expressar Experiências através de Regras**.



- Expressar Experiências através de Experiências.

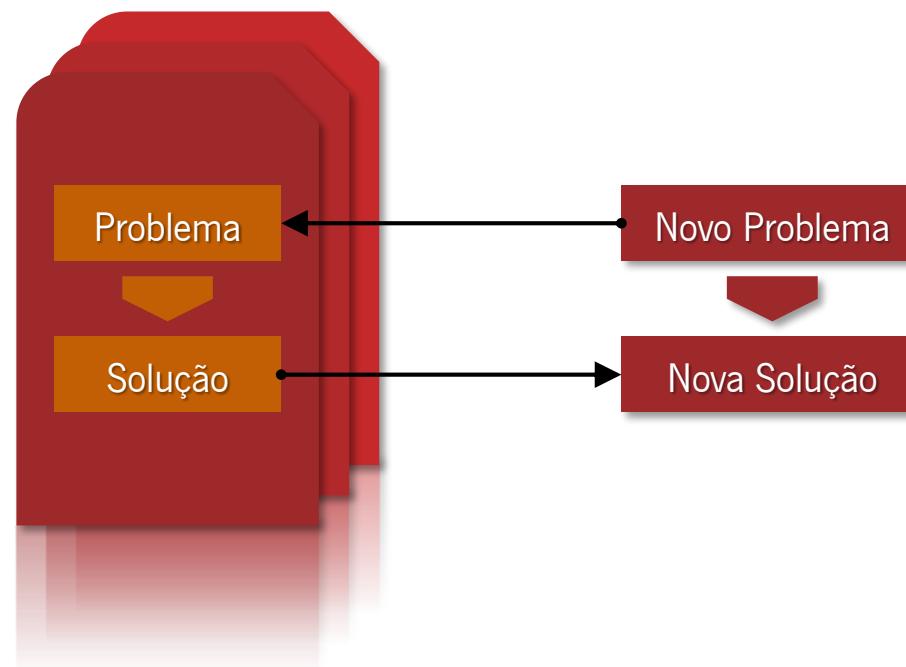


Experiência



Experiência

- A resolução de problemas passa pela adaptação de soluções existentes.



- Metodologia de resolução de problemas que consiste em **encontrar** e **justificar** a solução para um dado **problema** pela reconstituição de situações passadas **similares**, reutilizando ou adaptando conhecimento de tais situações.



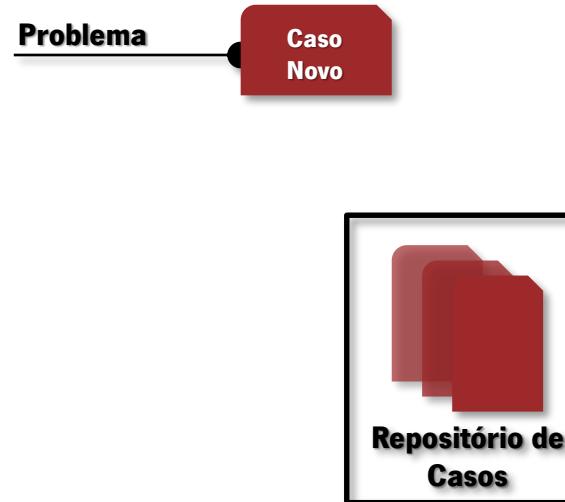
- **Casos são exemplos de situações passadas** que ocorreram no processo de resolução de um problema, consistindo na **Descrição do problema** em análise, na **Identificação da solução** para o problema e na correspondente **Justificação** da solução encontrada.



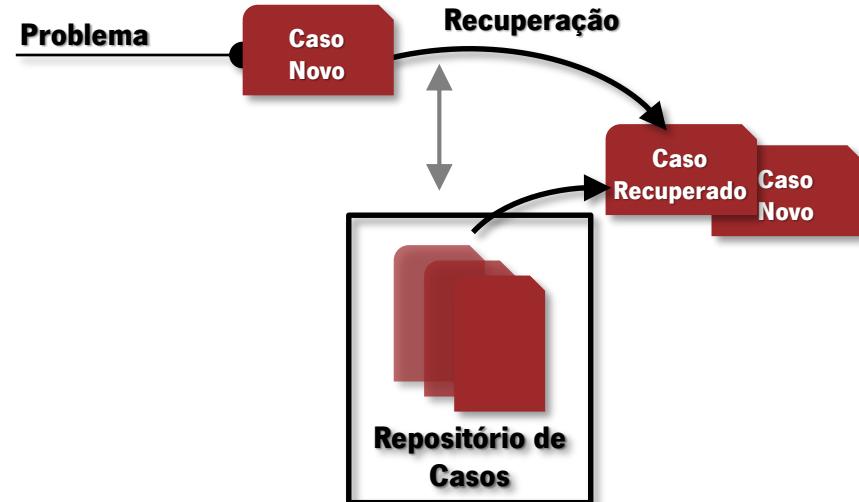
Áreas de aplicação

- Em domínios onde existe um grande volume de **conhecimento adquirido através de experiências**, mas eventualmente pobres em conhecimento ou mesmo conhecimento incompleto;
- Em áreas onde **especialistas** falam acerca do seu domínio através de exemplos;
- Quando existe um largo conjunto de **exceções a regras**;
- Se as tradicionais formas de representação de conhecimento não se encontram entre as mais apropriadas;
- **Sistemas onde a aprendizagem é fundamental.**

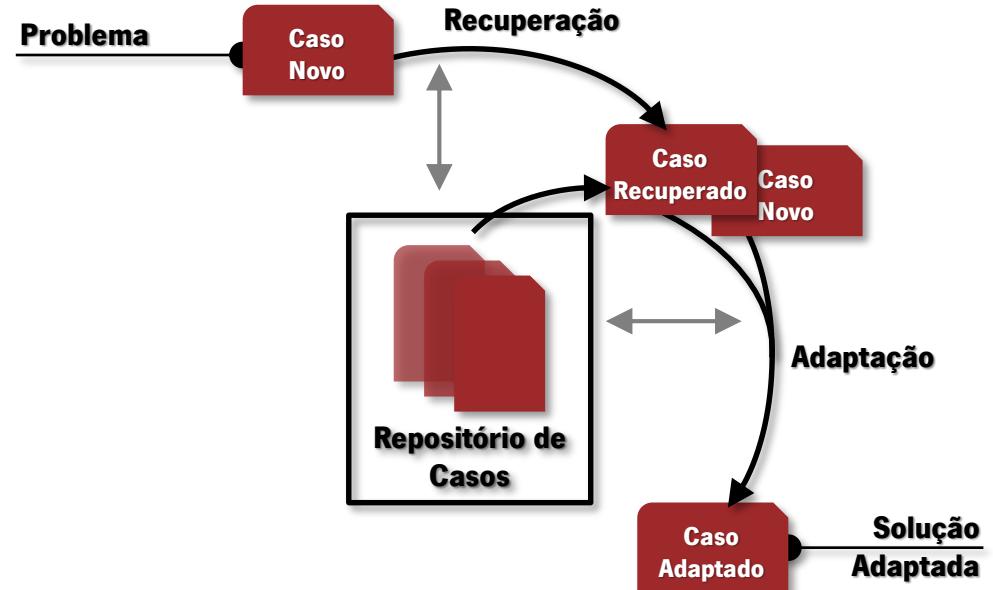
- 4 passos RBC (“os 4 R’s”):



- 4 passos RBC (“os 4 R’s”):
 - Recuperação (Retrieve);

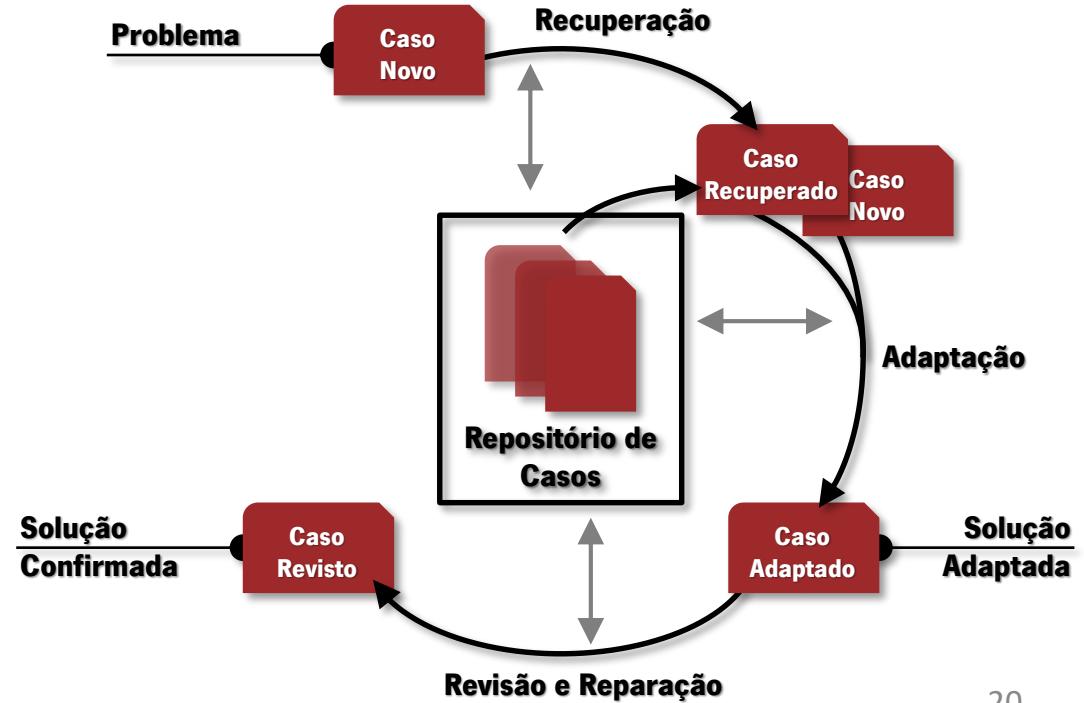


- 4 passos RBC (“os 4 R’s”):
 - Recuperação (Retrieve);
 - Adaptação (Reuse);



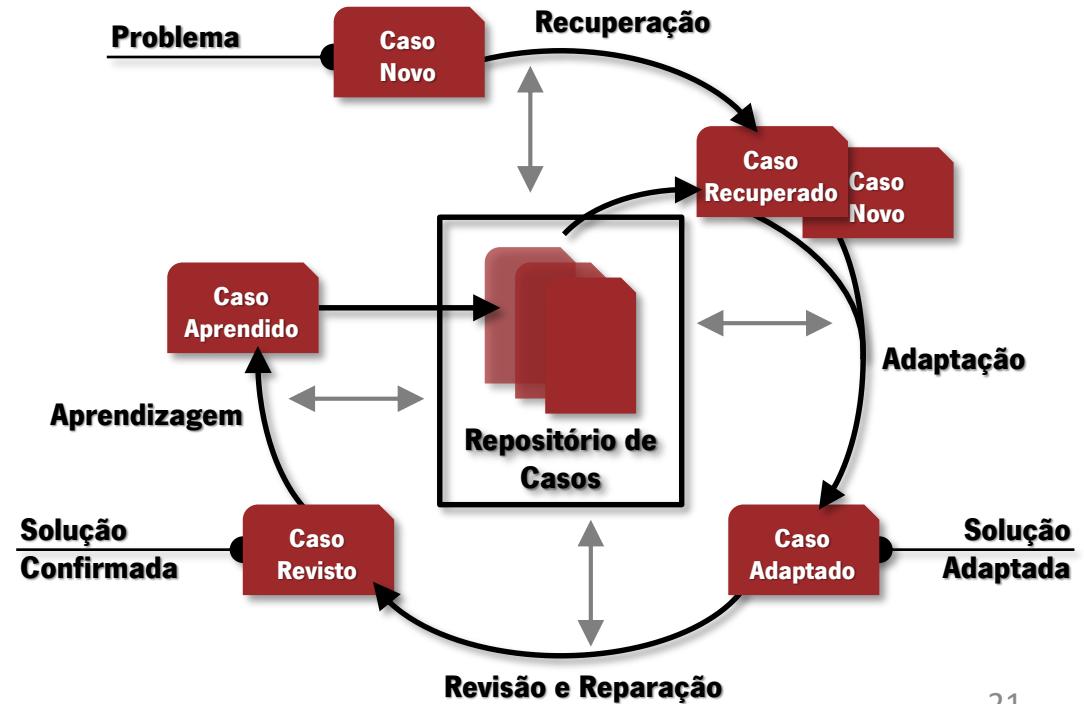
Modelo RBC

- 4 passos RBC (“os 4 R’s”):
 - Recuperação (Retrieve);
 - Adaptação (Reuse);
 - Revisão (Revise);



Modelo RBC

- 4 passos RBC (“os 4 R’s”):
 - Recuperação (Retrieve);
 - Adaptação (Reuse);
 - Revisão (Revise);
 - Aprendizagem (Retain).



Conteúdo de um Caso

- Um caso é um triplo $\langle P, S, J \rangle$ em que:
 - P é a descrição do **problema**;
 - S é a **solução** para o problema;
 - J é uma **justificação** para a solução.
- Um caso corresponde a um evento real ou a um processo, que pode ser limitado no tempo e no espaço.



Descrição do Problema

- Deve representar o estado do Problema:
 - Identificação da situação (características gerais);
 - Conjunto das características específicas do problema e as relações entre estas;
 - Objetivos a alcançar com a resolução do problema;
 - Restrições (invariantes) a observar na resolução do problema.



Solução do Problema

- Deve caracterizar-se por conter:
 - Descrição da solução propriamente dita;
 - Conjunto de passos usados para a sua realização;
 - Resultado da aplicação da solução.





- Deve caracterizar-se por conter:
 - Conjunto de decisões que foram tomadas na resolução do problema;
 - Métodos de resolução utilizados;
 - Explicações por parte do sistema computacional ou mesmo do utilizador quando é chamado a intervir.

Justificação da Solução





Conteúdo de um Caso: Extensões

- Extensões à representação de Casos:
 - outras soluções;
 - possíveis exceções (soluções não possíveis);
 - ligações a outros casos.



Escolha do Vocabulário

- A escolha do vocabulário utilizado na descrição dos casos pode ser examinado por duas vias:
 - **Funcional**, que passa pela análise do tipo de casos disponíveis e das tarefas que têm que ser suportadas pelo sistema (ou seja, escolhe-se aquilo que se tem);
 - **Recordação**, que passa pela captura das características e dos conceitos que são usados por especialistas, para descrever um determinado domínio ou universo de discurso.

Repositório de Casos

- Um dos fatores de sucesso dos sistemas de RBC consiste na faculdade de pesquisa e manuseamento fácil e rápido da Memória de Casos;
- Diversos modelos de estruturação do Repositório de Casos:
 - Memória não Hierarquizada (“não organizada”);
 - Modelo de Memória Dinâmico [Shank 82];
 - Modelo Baseado em Exemplos e Categorias [Bareiss et al 90].

A white, humanoid robot head with blue eyes and a teal neck and shoulder plate is positioned on the left side of the slide, partially cut off by the frame.

Exemplo: Uma Base de Casos

- Diagnóstico técnico de avarias de carros:
 - Observação de sintomas:
 - Ex. O motor não arranca;
 - Ex. Voltagem da bateria = 6.3 v.
 - Objetivo:
 - Encontrar a causa da avaria (Ex. bateria descarregada);
 - Estratégia de reparação (Ex. carregar a bateria).
- Base de casos de Diagnósticos:
 - Um caso descreve uma situação de diagnóstico e contém:
 - Descrição dos sintomas;
 - Descrição da avaria e da causa;
 - Descrição da estratégia de reparação.

- Descreve uma experiência particular de diagnóstico;
- Armazena as características mais significativas e os respetivos valores;
- **Não é uma regra**;
- Todos os casos devem ser independentes entre si.
- Caso 1:
 - Problema
 - Problema: Luzes da frente não funcionam
 - Carro: Ford Fiesta 1.8 D
 - Ano: 1995
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok
 - Solução
 - Diagnóstico: Fusível das luzes da frente com defeito
 - Reparação: Substituir fusível das luzes da frente

A Recuperação de Casos

- Tem como objetivo selecionar o(s) caso(s) que melhor combina(m) com a descrição do problema, desenvolvendo-se na forma:
 - identificação das características do problema em mãos;
 - procura dos casos que melhor combinam com a descrição do novo caso;
 - seleção.

Similaridades

- Na recuperação de casos a palavra chave é **similaridade**:
 - Preferências Heurísticas de Similaridade:
 - Medidas Heurísticas (adequado a modelos de memória hierarquizados);
 - Ex: A partilha do maior número de características.
 - Medidas de Similaridade:
 - Ex: Vizinho mais próximo.

Reutilização de Casos

- 
- O objetivo da **Reutilização de Casos** está no uso do conhecimento do(s) caso(s) passado(s), devolvido pela recuperação de casos, na resolução de um novo problema, que se processa na forma:
 - adaptação do conhecimento;
 - teste da solução;
 - reparação da solução.
 - Nesta fase, o feedback com o utilizador é essencial.
(a adaptação automática é “praticamente impossível”)



Exemplo: Uma base de Casos

- Caso 1:

- Problema

- Descrição: Luzes da frente não funcionam
 - Carro: Ford Fiesta 1.8 D
 - Ano: 1995
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Solução

- Diagnóstico: Fusível das luzes da frente com defeito
 - Reparação: Substituir fusível das luzes da frente

- Caso 2:

- Problema

- Descrição: Luzes da frente não funcionam
 - Carro: Ford Escort 1.8 TD
 - Ano: 1998
 - Voltagem da Bateria: 12.9v
 - Estado das Luzes: Aspecto estragado

- Solução

- Diagnóstico: Lâmpada com defeito
 - Reparação: Substituir lâmpada da frente

Resolução de um Problema

- Para todos os efeitos, um novo problema é um caso sem solução;
- Caso novo:
 - Problema
 - Descrição: Luz de travão não funciona
 - Carro: Ford Mondeo 2.0. TDDI
 - Ano: 2001
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok
 - Solução
 - ?
 - ?

Similaridades

- Caso novo:

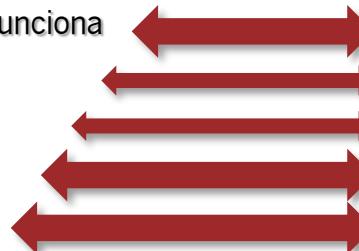
- Problema

- Descrição: Luz de travão não funciona
 - Carro: Ford Mondeo 2.0. TDDI
 - Ano: 2001
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Caso 1:

- Problema

- Descrição: Luzes da frente não funcionam
 - Carro: Ford Fiesta 1.8 D
 - Ano: 1995
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok



- Solução

- Diagnóstico: Fusível das luzes da frente com defeito
 - Reparação: Substituir fusível das luzes da frente

Similaridades

- Caso novo:

- Problema

- Descrição: Luz de travão não funciona
 - Carro: Ford Mondeo 2.0. TDDI
 - Ano: 2001
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Peso da característica de maior relevância: 6



- Peso da característica de menor relevância: 1



- Caso 1:

- Problema

- Descrição: Luzes da frente não funcionam
 - Carro: Ford Fiesta 1.8 D
 - Ano: 1995
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Solução

- Diagnóstico: Fusível das luzes da frente com defeito
 - Reparação: Substituir fusível das luzes da frente

- Similaridade(Caso novo, Caso 1) =
$$= 1/20 * (6*0,8 + 1*0,4 + 1*0,6 + 6*0,9 + 6*1,0)$$
$$= 0,86$$
- Similaridade(Caso novo, Caso 2) =
$$= 1/20 * (6*0,8 + 1*0,6 + 1*0,8 + 6*0,9 + 6*0,0)$$
$$= 0,585$$
- Similaridade(Caso novo, Caso 3) =
$$= \dots$$
$$= \dots$$
- ...



Similaridades

- Caso novo:

- Problema

- Descrição: Luz de travão não funciona
 - Carro: Ford Mondeo 2.0. TDDI
 - Ano: 2001
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Nova solução

- Diagnóstico: Fusível da luz de travão com defeito
 - Reparação: Substituir fusível da luz de travão

- Caso 1:

- Problema

- Descrição: Luzes da frente não funcionam
 - Carro: Ford Fiesta 1.8 D
 - Ano: 1995
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok

- Solução

- Diagnóstico: Fusível das luzes da frente com defeito
 - Reparação: Substituir fusível das luzes da frente

Adaptação

Técnicas de Adaptação

- 
- Adaptação nula (i.e., não necessita de adaptação);
 - Por substituição:
 - quando a modificação de atributos na solução é suficiente e não implica uma alteração da estrutura da solução (por exemplo, o ajuste de parâmetros, a instanciação);
 - Utilizando regras;
 - Por analogia:
 - consiste na transferência de conhecimento de soluções de problemas passados para o novo problema, usando-se o conhecimento transferido para construir a solução do novo problema.

- 
- O objetivo da aprendizagem é o de adquirir ou incorporar conhecimento resultante de experiências passadas, através da construção de um novo caso, modificando parte da base de conhecimento de um caso já existente, que se processa na forma:
 - seleção da informação que deve ser retida;
 - selecionar índices;
 - integração do caso na memória de casos.

- Caso 3:

- Problema
 - Descrição: Luz de travão não funciona
 - Carro: Ford Mondeo 2.0. TDDI
 - Ano: 2001
 - Voltagem da Bateria: 13.6v
 - Estado das Luzes: Ok
- Nova solução
 - Diagnóstico: Fusível da luz de travão com defeito
 - Reparação: Substituir fusível da luz de travão

- Se o diagnóstico está correto, deve-se armazenar o novo caso na base.



- Indução: forma de raciocínio em que se procura, a partir da verificação de alguns casos particulares, formular uma lei que explique todos os casos da mesma espécie (in Dicionário da Língua Portuguesa);
- RBC e Indução utilizam a mesma “matéria-prima” (casos);
- A Indução: cria uma abstração da memória de casos;
- O RBC: utiliza os casos passados durante a resolução de problemas;
- A Indução apresenta limitações na manipulação de informação incompleta durante o processo de generalização.

- 
- RBR: Raciocínio Baseado em Regras;
 - Descortinar as Regras mais relevantes para um dado domínio é um processo complexo, sendo as regras dificeis de articular e de avaliar;
 - Os Casos são facilmente estruturáveis e fáceis de encontrar;
 - O RBC pode evitar erros passados, além de ter a capacidade de gerar novas soluções através de um processo de generalização;
 - Num sistema RBR, as regras armazenam o conhecimento relativo à forma como se faz (em vez de descreverem a forma como funciona);
 - Num sistema RBC, usam-se situações anteriores como modelos para a conceção de explicações;
 - A tarefa de manutenção de sistemas de regras é complexo;
 - Sistemas de RBR são indicados quando é difícil obter ou não há disponibilidade de situações passadas.

- 
- As BDs devolvem informação armazenada;
 - O RBC pode adaptar a informação recuperada;

 - As BDs estão preparadas para encontrar a combinação entre a interrogação feita à base e a informação armazenada;
 - O objetivo do RBC é recuperar o caso mais similar;

 - As BDs operam sem qualquer conhecimento do problema, limitando-se a fornecer mecanismos de pesquisa e recuperação de informação;
 - Os sistemas de RBC usam o conhecimento acerca do problema para construir os mecanismos de recuperação de casos.

- 
- As RNAs e os RBC utilizam casos; contudo, nas RNAs, os casos servem apenas para o treino e teste da rede;
 - As RNAs têm melhor desempenho que os RBC em domínios com informação incompleta ou mesmo de falta total de informação;
 - Nas RNAs o utilizador não pode julgar (ou avaliar) as decisões da rede; não existem justificações ou explicações para as decisões.

Combinando com outros métodos

- Indução;
- Sistemas Baseados em Regras;
- Como componente de um sistema multi-estratégia de aprendizagem;
- Em associação com outros métodos de raciocínio para melhorar e guiar o processo de aprendizagem do RBC.

- 
- Resolução de Problemas:
 - Reutilização de soluções:
 - CHEF - receitas chinesas [Hammond 86].
 - Reutilização do processo de derivação de soluções:
 - PRODIGY/Analogy - problemas de logística [Veloso 94].
 - Classificação/Interpretação:
 - Classificação de situações ou objetos:
 - PROTOS - [Bareiss et al 89];
 - Direito:
 - HYPO - [Ashley and Rissland 88]
 - Ensino Assistido:
 - Sistemas interativos que fornecem informações:
 - No ensino [Shank 95].

Aplicações Comerciais

- As aplicações comercializadas são RBC Shells e fornecem:
 - editores de memória de casos;
 - mecanismos de indexação e recuperação de casos;
 - mecanismos limitados de adaptação (ex: ajuste de parâmetros baseado em regras).
- Exemplos de Shells disponíveis no mercado:
 - ReMind (Cognitive Systems inc., USA);
 - CBR-Express (Inference, USA);
 - KATE-Tools (Acknosoft, France&USA);
 - CBR Works (Tecinno Gmbh);
 - MyCBR (DFKI Gmbh);
 - ...

Futuras Direções

- Criação de componentes inteligentes de RBC;
- Integração com outros sistemas;
- Aperfeiçoamento da adaptação automática de casos;
- Utilização do modelo de RBC no comércio eletrónico.

- “[...] transferring knowledge from past problem solving episodes to new problems that share significant aspects with corresponding past experience and using the transferred knowledge to construct solutions to new problems.”

Carbonell, 1986



“If we learn from our mistakes, shouldn’t I try to make as many mistakes as possible?”

- 
- Aceder ao projeto My-CBR:
 - <http://www.mycbr-project.net>
 - Descarregar a aplicação My-CBR:
 - <http://www.mycbr-project.net/download.html>
 - Descarregar o tutorial (slides):
 - <http://www.mycbr-project.net/tutorials.html>
 - Em [[Downloads](#)], da secção “Sample projects”, descarregar os dois ficheiros do projeto “Cars project” (.zip e .csv);
 - Seguir o tutorial a partir da página 90.

- 
- Instalar a aplicação CBR-Works;
 - Usar a ferramenta Case Navigator:
 - Executar a Query, com os seguintes atributos:
 - Color: yellow
 - Manufacturer: Audi
 - Model: A4 1.6
 - Year: 1997
 - Executar a recuperação de casos;
 - Verificar a possibilidade de proceder a algum tipo de adaptação;
 - Observar a organização e estruturação da memória de casos.

- 
- Desenvolva um repositório de casos para uma situação de diagnóstico médico.
 - No modelo de representação de casos deverá ter em conta que cada componente contenha a informação seguinte:
 - descrição do problema, os sintomas de uma doença, bem como a informação relacionada com o paciente;
 - solução do problema, o resultado do diagnóstico efetuado;
 - Justificação, terapêutica aplicada e eventuais contra-indicações.
 - Elaborar uma pequena aplicação que faça o armazenamento, numa estrutura não hierarquizada, dos casos.
 - Implemente a fase de recuperação de casos (1^a fase do modelo apresentado), usando como medida de similaridade a partilha do maior número de sintomas entre o novo problema e casos passados.

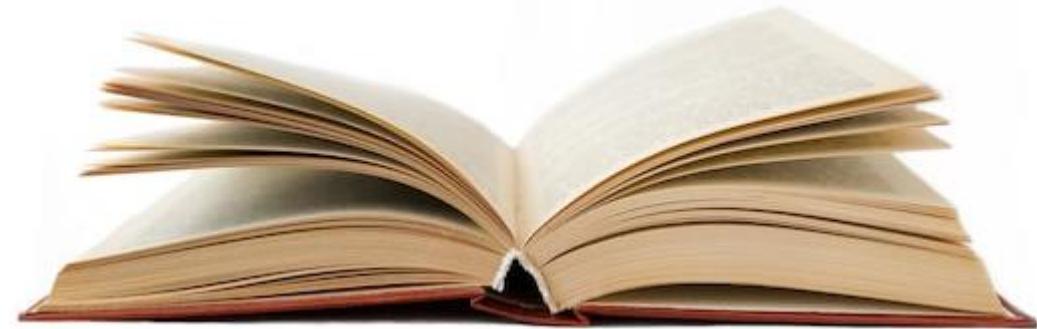
- 
- Desenvolva um repositório de casos para o conjunto de dados “meteo.csv”:
 - Criar novo projeto (“Meteo”);
 - Definir o conceito raiz (“Praticar”);
 - Importar o ficheiro com os casos (meteo.csv);
 - Definir regras de cálculo de similaridade para cada atributo;
 - Experimentar a recuperação de alguns casos:
 - definir fórmula de cálculo de similaridade global;
 - experimentar diversos cenários de cálculo da similaridade global;
 - Realizar a retenção de novo caso resolvido.

Exercício V

- 
- Desenvolva um repositório de casos para o conjunto de dados “creditset.csv”:
 - Criar novo projeto (“Crédito”);
 - Definir o conceito raiz (“Empréstimo”);
 - Importar o ficheiro com os casos (creditset.csv);
 - Definir regras de cálculo de similaridade para cada atributo:
 - definir duas regras de cálculo de similaridade para cada atributo;
 - Experimentar a recuperação de alguns casos:
 - definir fórmula de cálculo de similaridade global;
 - experimentar diversos cenários de cálculo da similaridade global;
 - Realizar a retenção de novo caso resolvido.

Referências bibliográficas

- Aamodt A., Plaza E., "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", in AI Communications, Vol. 7, N° 1, pages 39-59, 1994.
- Leake D., "Case-Based Reasoning - Experience, Lessons and Future Directions", in The MIT Press, 1996.
- Novais P., Neves J., "Raciocínio Baseado em Casos", Universidade do Minho, 1998.
- Sun Z. and Finnie G., "Intelligent Techniques in E-Commerce: A Case-based Reasoning Perspective", Springer, ISBN: 978-3-540-20518-0, 2004.



Universidade do Minho

Escola de Engenharia

Departamento de Informática

Aprendizagem e Extração de Conhecimento

Perfil Sistemas Inteligentes @ MiEI/4º – 1º Semestre

Cesar Analide, Filipe Gonçalves