Raciocínio Baseado em Casos - RBC

RUDIMAR LUÍS SCARANTO DAZZI

Apresentação

- Definição
- Histórico
- Representação de conhecimento
- O que é um caso?
- Ciclo do RBC
- Aplicações de exemplo
- Conclusão

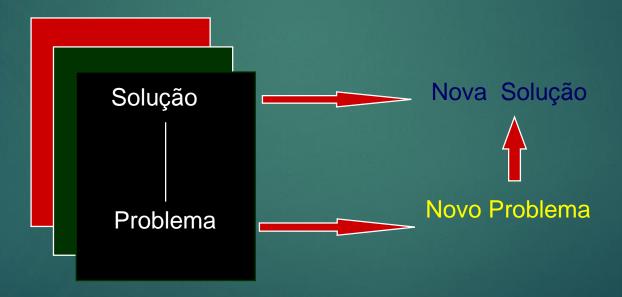
- CBR: Case-based Reasoning (RBC: Raciocínio Baseado em Casos)
- Área da Inteligência Artificial
- Corrente Simbólica SBC
 - Automatizar o processo de solução de problemas através do conhecimento específico
- Baseado no modelo de armazenamento de informações na memória humana
- Solução de novos problemas através de experiências passadas

- Domínio particular
- Problemas tendem a ser parecidos e com poucas alterações
- Visão oposta: MBR Model-based reasoning
 - Modelo genérico de representação do domínio

"Um sistema de RBC resolve problemas por adaptar soluções que foram utilizadas para resolver problemas anteriores." Riesbeck e Schank - 1989

.... transforma conhecimento a partir da solução de um episódio passado para um novo problema que tem aspectos significantes com experiências passadas correspondentes e usando conhecimento transferido para construir soluções para novos problemas." (Carbonell, 1986)

▶ RBC - é uma metodologia de solução de problemas por adaptar soluções anteriores de problemas similares.



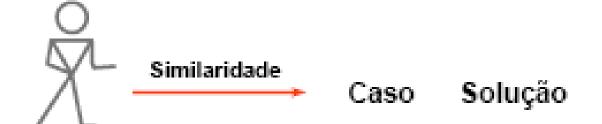
Raciocínio Baseado em Casos

Experiência

Estudante de Médico Especialista

Medicina

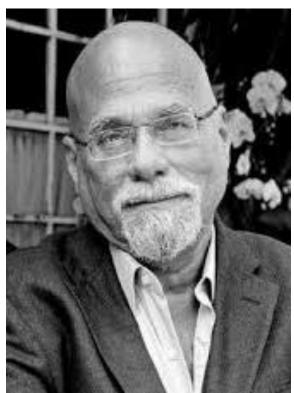
Regras Diagnósticos Compreensão Princípios de Pacientes Indexação



Histórico

- Início anos 80, pesquisas de Roger Schank
- Conhecimento registrado em scripts
- Áreas da Psicologia e Computação
- Janet Kolodner: primeira aplicação





O que é o Raciocínio Baseado em Casos ?

Um sistema semelhante a uma base de casos (armazenamento indexado de casos) passados que são reproduzidos em um caso atual.

A indexação e representação de casos facilita a recuperação de casos relevantes e a comparação deles com o problema atual.

Aplica informação em casos recuperados para analisar ou resolver o problema.

Representação de conhecimento

- O tratamento de conhecimento em mecanismos artificiais como os computadores é uma tarefa difícil.
- Para que se possa manipular este conhecimento é necessário que se consiga forma de representá-lo.
- A representação do conhecimento é uma das áreas mais ativas da IA, e a que envolve os maiores desafios

Representação de conhecimento

- Em um sistema RBC a representação de conhecimento é feita principalmente em forma de casos que descrevem experiências concretas.
- Mesmo que o conhecimento do especialista esteja em alguns momentos representado em RBC, são os casos que contém o conhecimento utilizado para resolução do problema

Caso

- Um caso é um pedaço de conhecimento que está de acordo com determinado contexto e representando uma experiência de onde se aprende lições fundamentais para atingir os objetivos do raciocinador.
- A experiência contida no caso deve estar bem descrita, e de acordo com seu contexto.

Caso

Exemplo simples de um caso com a experiência adquirida

Nome: Paulo Rocha

Nascimento: 20/05/1964

Endereço: Av. Carlos Gomes, 45

Profissão: Analista de Sistemas

Salário mensal: R\$ 3.000,00

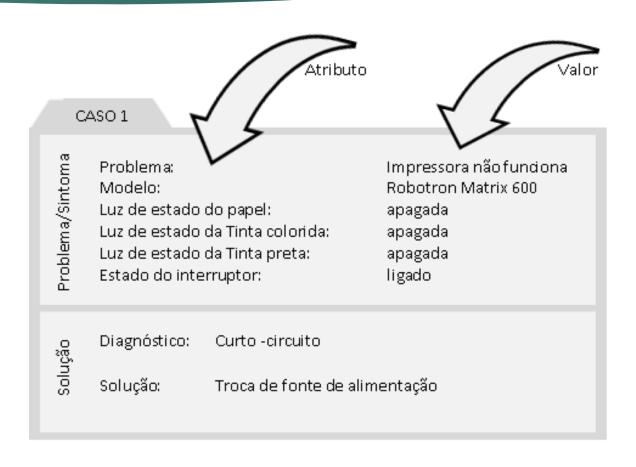
Estado civil: Casado

Dependentes: 0

Cartão de crédito: Visa

Empréstimo: R\$ 20.000,00

Concedido: SIM



Base de casos

- Consiste em um conjunto de casos e procedimentos para acessar estes casos
- A base de casos serve para que se possam reutilizar os casos organizados e armazenados nela
- Contém geralmente experiências positivas que descrevem estratégias de solução para serem reutilizadas
- Contém experiências negativas que expressam tentativas frustradas de se solucionar um problema

Base de casos

Exemplo da representação de casos em uma base casos em uma base

CASO 1 Problema/Sintoma Problema: Impressora não funciona Robotron Matrix 600 Modelo: Luz de estado do papel: apagada Luz de estado da Tinta colorida: apagada Luz de estado da Tinta preta: apagada Diagnóstico: Solução Curto -circuito Solução: Troca de fonte de alimentação

CASO 2 Problema/Sintoma Problema: Não imprime em preto Modelo: Robotron Matrix 800 Luz de estado do papel: apagada Luz de estado da Tinta colorida: apagada Luz de estado da Tinta preta: aceso Diagnóstico: Solução Curto -circuito Solução: Troca de fonte de alimentação

Ciclo do RBC

- Aquisição de casos
- Representação de casos
- Indexação
- Recuperação
- Adaptação
- Solução/Aprendizado

- Processo mais complexo no desenvolvimento de uma aplicação, já que são várias as formas de se realizar a aquisição das informações
 - Os dados podem não estar disponíveis em uma fonte externa
 - Estão semi-disponíveis em uma fonte externa
 - Existe o registro das informações porém, com erros
 - Os dados estão registrados de forma correta e disponíveis

- Os dados podem não estar disponíveis em uma fonte externa
 - Técnica de RBC é inviável
 - Esforço de adquirir casos é o mesmo de construir um modelo e modelos são mais eficientes

- Os casos estão semi-disponíveis em uma fonte externa
 - Situação mais comum de ser encontrada
 - Aquisição de conhecimento com o tomador de decisão pode complementar a informação não-disponível

- Existe o registro das informações porém, com erros
 - Situação muito comum
 - Os casos selecionados para realmente integrar a base de casos devem ser verificados e validados pelo tomador de decisão

- Os casos estão registrados de forma correta e disponíveis
 - Situação muito incomum
 - Apenas os casos duplicados ou não representativos devem ser excluídos da base

Vantagens da Aquisição de Conhecimento por Casos

- Construção de um protótipo antes de obter a completa estruturação do domínio
- Pode ser realizada mesmo em domínio pouco estruturados ou com base teórica mal definida
- Casos são uma boa amostragem dos tipos de problemas que o sistema deve resolver
- Diminui a necessidade de entrevistas
- Encapsulamento do conhecimento

Representação de casos

- Os casos representam um conhecimento especifico que está diretamente relacionado a um determinado contexto, e armazena um conhecimento de nível operacional.
- Casos armazenados devem apresentar diferenças dos demais para que seja acrescentado algum conhecimento útil para o sistema.
- Características que descrevem o caso devem apontar para uma solução ao problema inicial

Representação de casos

 Exemplo da representação de um caso da área jurídica DJ: 5.555 DATA: 17/05/90 PÁG: 08

Apelação criminal n. 55.824, de Hipoteticópolis da Serra. Relator: Des. Antônio Empederneiras.

APELAÇÃO CRIMINAL. PEDIDO DE DESISTÊNCIA. HOMOLOGAÇÃO.

Vistos, relatados e discutidos estes autos de apelação criminal n. 55.824, da comarca de Hipoteticópolis da Serra, em que é apelante Cecolino Cabresto, sendo apelada a Justiça, por seu promotor:

ACIRDAM, em Primeira Câmara Criminal, à unanimidade, homologar a desistência requerida.

Custas de lei.

Trata-se de pedido de desistência do recurso interposto por defensor em favor de CECOLINO CABRESTO que na comarca de Hipoteticópolis da Serra foi condenado a pena de 12 (doze) anos e 6 (seis) meses de redusão, por infração ao art. 121, s 2º, IV c/c art. 14, ambos do Código Repressivo. Presente os pressupostos que autorizam o acolhimento da pretensão, homologase o pedido de desistência. Presidiu o julgamento o Exmo. Srs Des. Rigorosíssimo Praga e participaram do mesmo, com votos vencedores, os Exmos Srs. Des. Zélio Botaolho e Mad Mortis.

Capitalópolis, 01 de abril de 1990.

Ludovico da Várzea
Presidente p/ o acórdão
Rigorosíssimo Praga
Relator
Rigorosíssimo Praga
Procurador de Justica

CASO 1

Número do acórdão: 55824

Data da publicação:17/0 5/90

Localização: Hipoteticópolis da Serra

Tipo de Recurso: Apelação criminal

Relator: Rigorosíssimo Praga

Resultado: Concedido

Tipificação

Tipo geral: Homicídio **Modalidade:** Doloso

Qualificação: Homicídio Qualificado

Tentativa: Sim Co-autoria: Não

Expressões indicativas: Pedido de desistência do recurso; Homologação do pedido de desistência; Presentes

os pressupostos

DJ: 5.555 DATA: 17/05/90 PÁG: 08

Apelação criminal n. 55.824, de Hipoteticópolis da Serra.

Relator: Des. Antônio Empederneiras.

APELAÇÃO CRIMINAL. PEDIDO DE DESISTÊNCIA. HOMOLOGAÇÃO.

Vistos, relatados e discutidos estes autos de apelação criminal n. 55.824, da comarca de Hipoteticópolis da Serra, em que é apelante Cecolino Cabresto, sendo apelada a Justiça, por seu promotor:

ACIRDAM, em Primeira Câmara Criminal, à unanimidade, homologar a desistência requerida.

Custas de lei.

Trata-se de pedido de desistência do recurso interposto por defensor em favor de CECOLINO CABRESTO que na comarca de Hipoteticópolis da Serra foi condenado a pena de 12 (doze) anos e 6 (seis) meses de redusão, por infração ao art. 121, s 2º, IV c/c art. 14, ambos do Código Repressivo. Presente os pressupostos que autorizam o acolhimento da pretensão, homologase o pedido de desistência. Presidiu o julgamento o Exmo. Srs Des. Rigorosíssimo Praga e participaram do mesmo, com votos vencedores, os Exmos Srs. Des. Zélio Botaolho e Mad Mortis.
Capitalópolis, 01 de abril de 1990.

Ludovico da Várzea Presidente p/ o acórdão Rigorosíssimo Praga Relator Rigorosíssimo Praga Procurador de Justiça

CASO 1

Número do acórdão: 55824 Data da publicação:17/0 5/90

Localização: Hipoteticópolis da Serra **Tipo de Recurso:** Apelação criminal

Relator: Rigorosíssimo Praga

Resultado: Concedido

<u>Tipificação</u>

Tipo geral: Homicídio **Modalidade:** Doloso

Qualificação: Homicídio Qualificado

Tentativa: Sim Co-autoria: Não

Expressões indicativas: Pedido de desistência do recurso; Homologação do pedido de desistência; Presentes

os pressupostos

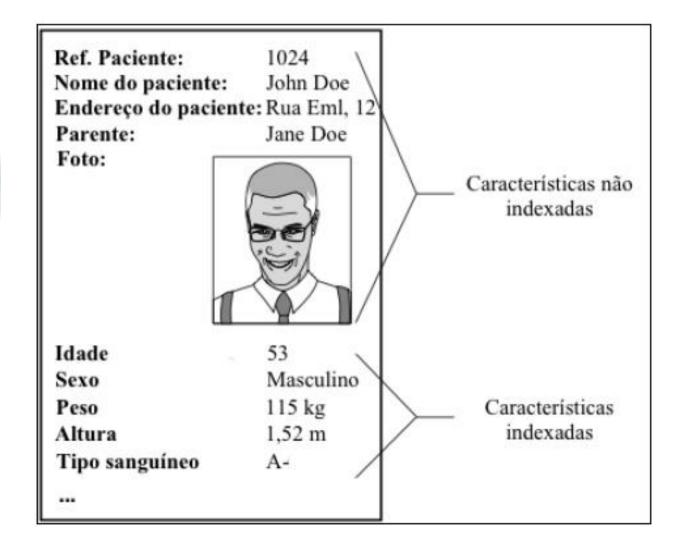
Representação de casos

- Abordam também que a representação de conhecimento dos casos deve ser bem definida para formular o conhecimento da solução do problema
- Podem seguir uma série de objetivos como:
 - Definir as entidades do domínio da aplicação;
 - Descrever as dependências e relacionamentos das entidades em questão;
 - Manutenção da estrutura de domínio da aplicação;
 - Prevenir representação redundante de conhecimento; e
 - Suportar definição de similaridade e adaptação de soluções.

- A indexação na recuperação de casos é essencial, pois orienta a avaliação da similaridade.
- Descritores são utilizados para os índices para verificar o que faz um caso ser similar ao outro, representando a relevância entre os casos.
- A indexação vai determinar o que precisa ser comparado entre os casos para avaliar sua similaridade, para que sejam recuperados os casos mais úteis interpretando o novo caso

A indexação é que determina o que deve ser comparado entre os casos, determinando sua similaridade.

Isso faz com que o RBC, além de utilizar os índices para melhorar a rapidez na recuperação, também utilize para fazer atribuição de similaridade entre os casos com eficiência.



- Um critério importante da indexação, quando atribuídos os índices, é assegurar que o caso será acessado sempre que apropriado.
- A similaridade baseada nestes índices é que vai predizer a utilidade do caso em relação a uma meta específica de recuperação

- Portanto a definição de índices deve ser feita cuidadosamente
- Características superficiais e menos relevantes são facilmente extraídas, mas são também menos úteis que índices mais complexos.

Métodos de Indexação - Manual

- 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 um checklist
- Indexar desta forma é praticamente um trabalho de aquisição de conhecimento

Métodos de Indexação -Automático

- Uma análise matemática do comportamento da base sugere os melhores índices
- Indução e árvores de decisão
- Indexação de casos através dos cálculos de sensibilidade e especificidade

A recuperação de casos é utilizada para, a partir da descrição de um problema, encontrar um conjunto de casos similares ao problema analisado que seja útil para identificar a sua solução.

- A partir de um problema de entrada, a recuperação faz uma busca na memória de casos.
- Esta busca é feita utilizando algoritmos que selecionam casos com uma similaridade determinada em relação ao problema de entrada,
- Resulta em um novo caso da base a ser reutilizado.

- Divide tarefas da descrição do problema até o encontro do melhor caso.
- A tarefa divide-se em:
 - Identificação das características: informa ao sistema características do caso atual;
 - Unificação inicial: possíveis candidatos são recuperados;
 - Busca: seleciona qual o melhor candidato entre os casos recuperados no casamento inicial; e
 - Seleção: os casos são ordenados de acordo com a métrica de classificação, de acordo com o caso que possui similaridade mais forte com o problema conhecido

- Além disso alguns outros fatores são importantes na etapa de recuperação:
 - Eficiência: velocidade do sistema ao recuperar os casos;
 - Precisão: grau de utilidade dos casos que podem ser utilizados para atingir o objetivo proposto; e
 - Flexibilidade: grau de recuperação em caso de raciocínios inesperados.

Recuperação

- No vizinho mais próximo a similaridade entre um caso novo e um já existente é determinada para cada atributo.
- A medida encontrada deve ser multiplicada pelo peso e calculada a somatória de todos os atributos
- Com isto se estabelece à medida de similaridade entre os casos
 - Métodos mais utilizados são:
 - Algoritmo de Vizinhança (Nearest Neighbour)
 - Algoritmo de Indução (ID3, ID4, ID5)
 - Algoritmo de Indução guiada por conhecimento
 - Algoritmo de Recuperação de Padrões

- A similaridade é um ponto crucial do RBC, pois é uma etapa que torna o processo de raciocínio viável, avaliando-se a similaridade dos casos de entrada com os casos candidatos (base de casos).
- A similaridade entre os casos se baseia em características que representam o conteúdo e contexto de sua experiência.
- Em RBC a avaliação de similaridade é feita através da comparação de dois casos para avaliar o relacionamento entre eles

Esta avaliação é feita no nível de atributos onde valores são associados comparando valores numéricos e alfanuméricos

A utilidade de um caso é parte integrante do conceito de

similaridade a se adotar

Nome: Paulo Rocha

Nascimento: 20.05.64

Endereço: Av. Carlos Gomes, POA

Profissão: Analista de sistemas

Salário mensal: \$3000 Estado civil: solteiro

Dependentes: 0

Cartão crédito: Visa

Nascimento: 23.06.60

Profissão: Analista de sistemas

Salário mensal: \$3000 Estado civil: solteiro

Dependentes: 0 Cartão crédito: Visa

Empréstimo solicitado: \$20000

Empréstimo concedido: sim

- Um caso pode ser útil para um novo problema quando:
 - Permita a solução do problema atual;
 - Evita que um erro anterior seja repetido;
 - Permita uma solução rápida e eficiente do problema;
 - Ofereça a melhor solução para o problema baseado em critérios de otimização; e
 - Ofereça ao usuário uma solução na qual ele possa compreender a lógica utilizada.

- A avaliação de similaridade é dividida em dois grupos: a similaridade sintática e a similaridade semântica.
 - A similaridade sintática é mais simples, sendo sua comparação principal feita por semelhanças na sintaxe.
 - A similaridade semântica é mais profunda e os casos são comparados baseados no seu significado.

Cálculo de Similaridade

- Um caso possui diversos atributos de diferentes tipos
- Cada atributo pode ter uma função específica de cálculo de similaridade

Função de Peso do Número similaridade atributo de casos $\int w_i \times sim(f_{i'}^l f_i^R)$ i=1Valor do atributo i para o caso problema Valor do atributo i Soma dos para o caso pesos para i da base 52

Algoritmo de Vizinhança

Similaridade de Tipos Numéricos

- Para um tipo inteiro com distribuição linear e uniforme de valores (todos os valores se distribuem por todo o intervalo válido)
- Exemplo para tipo numérico:
- a1 = 40
- a2 = 80
- sim(a1,a2) = 1 |a2 a1| / (max min)
- Supondo que min = 0 e max = 100 :
- sim(40,80) = 1 |80 40|/(100 0) = 0.6

Similaridade de Tipos Numéricos

- Considera uma distribuição uniforme (linear) para os valores dos atributos
- Distribuições mais complexas devem usar funções que avaliam a distribuição

Similaridade de Tipos Simbólicos

- São definidas enumerações
- Exemplo:
 - Cores = {Branco, Amarelo, Vermelho, Marrom, Preto}
 - A1 = Branco
 - A2 = Amarelo

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

Similaridade de Tipos Simbólicos

- Opção2: Ordenar as distâncias entre os valores de forma uniforme
- a1 = Branco
- a2 = Amarelo
- \rightarrow sim(a1,a2) = 1 |0,25 0| / 1 = 0,75

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

Similaridade de Tipos Simbólicos

Opção 3: Criar uma matriz com as similaridades entre os valores

	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

sim(a1,a2) = 0.8

Adaptação

- Procura por diferenças salientes entre as duas descrições e aplica regras para compensá-las
- As regras de adaptação podem ser automáticas ou manuais
- Na maioria das aplicações desenvolvidas, a adaptação é realizada pelo usuário do sistema, já que o uso da adaptação automática, se realizada de forma errada, pode comprometer a confiabilidade

Adaptação: Ajuste de Parâmetros

Novo Caso: Clementino bateu em Raimundo. Raimundo ficou relativamente machucado. Raimundo revidou, batendo em Clementino muitas vezes. Clementino cai, levanta-se e continua a bater em Raimundo, quebrando-lhe o nariz. Raimundo tira uma faca e ameaça Clementino, que continua a brigar e acaba ferido e morrendo.

Adaptação: Ajuste de Parâmetros

- Caso Antigo: Benedito bateu em Reinaldo. Reinaldo ficou relativamente machucado. Reinaldo revidou, batendo em Benedito muitas vezes. Benedito cai, levanta-se e continua a bater em Reinaldo, quebrando-lhe as costelas. Reinaldo mata Benedito com vários tiros. Reinaldo foi condenado a 25 anos de prisão.
- Regra modificando a sentença de Reinaldo para adaptá-la à nova situação com atenuantes.

Adaptação

- Pode-se fazer adaptações de várias formas:
 - Incluindo um novo comportamento na solução recuperada;
 - Eliminando um comportamento da solução recuperada; e
 - Substituindo parte de um comportamento



Adaptação: Substituição baseada em casos

Exemplo: Um problema de preparação de um menu, temos um convidado que é vegetariano e temos que substituir uma lasanha a bolognesa por outro prato. Ao invés de procurar por outros pratos na classe de massa que não tenham a restrição "sem carne", o conjunto de casos armazenados na base que incluem "massas" é percorrido, à busca de uma alternativa para lasanha.

Adaptação: Métodos de transformação

- Exemplo: Problema da preparação do menu para um grupo de convidados. Como na situação anterior, já tínhamos selecionado o caso com o menu de lasanha bolognesa. Mas lasanha vai contra a restrição "vegetariano". Mas aqui, ao invés de substituir a lasanha por outro prato, nós vamos tentar transformar a lasanha em "lasanha vegetariana".
- Como carne é um ingrediente secundário numa lasanha, podemos simplesmente remover carne do prato através de uma heurística chamada: remover ingrediente secundário.
- Outra solução seria substituir leite por outro produto que cumpriria a mesma função no prato, usando a heurística de substituição chamada "substituir item", por carne de soja por exemplo

Adaptação

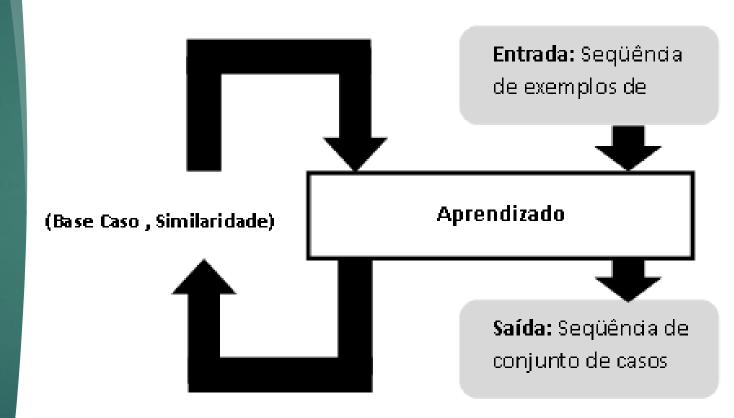
- Algumas questões centrais para estas adaptações são:
 - Quais os aspectos da situação em um caso que devem ser adaptados;
 - Que medidas são razoáveis para se modificar um caso;
 - Quais métodos de adaptação são aplicáveis na modificação do caso; e
 - Como controlar a adaptação para que não se perca o controle sobre ela.

Aprendizado/Solução

- O fator principal para aprendizagem em um sistema
 RBC acontece pela acumulação de novas experiências
 na base de casos e problemas bem indexados.
- O sistema se torna mais eficiente e apto a aprender

Aprendizado/Solução

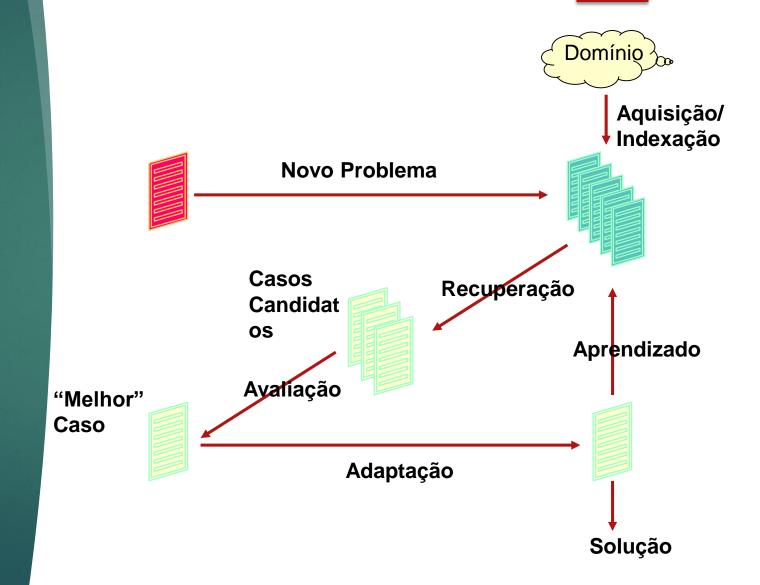
A aprendizagem pode ser vista como um processo de melhora de um sistema RBC



Aprendizado/Solução

- Quanto mais casos estiverem contidos na base de casos, que tenham o mesmo conjunto de atributos que definem a similaridade, maior a contribuição para solução de um novo caso.
- Aproveitar experiências passadas para antecipar e evitar erros melhora muito a qualidade da solução

Ciclo do RBC



Vantagens de RBC

- Facilita a aquisição de conhecimento
- Amostragem antecipada do tipo de problemas
- Reuso de conhecimento armazenado em bancos de dados e outras fontes
- Encapsulamento da descrição do conhecimento com a solução aplicada
- Aprendizagem automática de novos casos

Desvantagens de RBC

- Obter casos disponíveis e confiáveis
- Não cobrem todo o domínio
- Recuperação por similaridade é mais custosa do que analise de regras.
- Não existem bons algoritmos de adaptação
- Espaço de armazenamento maior do que modelos

Conclusão

- Pode ser aplicado em qualquer área
- Pode ser utilizado em conjunto com outras tecnologias de IA
- Direcionado para o gerenciamento de informação
- Automação de tarefas repetitivas

Bibliografia

- ABEL, Mara.; CASTILHO, J.M.V Raciocínio Baseado em Casos, PPGC-UFRGS, 1996.
- Utilização de RBC para determinação de aparelhos auditivos (artigo).
 Maio/2001
- ► LORENZI, F.; ABEL M. Aplicando raciocínio baseados em casos na investigação de irregularidades nas internações hospitalares, Il Congresso Brasileiro de Computação, 2002.
- MARTINS, Rafael B. Sistema em RBC para apoio a decisão nos processos judiciais da área de família. Itajaí, 2010. CTTMAT, UNIVALI, Itajaí, 2010.



http://www.ai-cbr.com

http://www.inf.ufrgs.br/gpesquisa/bdi