

Sistemas Inteligentes Adaptativos

Programa de Ingeniería de Sistemas

Tema: Razonamiento Basado en Casos - RBC





Razonamiento Basado en Casos

Definición

El Razonamiento Basado en Casos (RBC ó CBR) es una técnica de la AI que intenta llegar a la solución de nuevos problemas de forma similar como lo hacen los seres humanos utilizando la experiencia acumulada hasta el momento en acontecimientos similares (Fowler 2000; Rossillea et al., 2005). Un nuevo problema se compara con los casos almacenados previamente en la base de casos (Memoria de Casos) y se recuperan uno o varios casos.

Posteriormente, se utiliza y evalúa una solución sugerida, por los casos que han sido seleccionados con anterioridad, para tratar de aplicarlos al problema actual (Delgado 2003; Reyes & Sison, 2002; Salcedo 2002; Althoff 2001; Manjares 2001; Riebech 1999; Cuenca 1998; Rich 1994).



Razonamiento Basado en Casos

Características

- ✿ El RBC resuelve nuevos problemas usando o adaptando soluciones que fueron usadas para resolver anteriores problemas.
- ✿ El RBC ofrece un paradigma de razonamiento que es similar a la manera en que muchas personas resuelven problemas rutinariamente.



Razonamiento Basado en Casos

Anexo: RBC como modelo del razonamiento humano

Queremos decidir el menú de una cena. Entre los invitados hay vegetarianos y Ana, que es alérgica a los productos lácteos. Como es la temporada del tomate, queremos usarlos en el plato principal.

- ✿ Una vez que tenía vegetarianos a cenar preparé pastel de tomate como plato principal y tuve mucho éxito.
- ✿ Pero este plato contiene queso y Ana no lo puede comer.
- ✿ Sin embargo, recuerdo que en otras ocasiones he adaptado recetas para Ana, sustituyendo el queso por tofu. Podría hacerlo así, pero no sé como resultará el pastel de esta forma.
- ✿ Decidimos descartar la tarta de tomate.
- ✿ Estaría bien hacer pescado a la parrilla. Aunque la última vez descubrí que a Elena no le gusta el pescado y tuve que hacer unos perritos calientes a última hora.
- ✿ Sin embargo, recuerdo que en una ocasión Elena comió pez espada. Quizá sea de esas personas a las que les gusta el pescado que parece carne.
- ✿ Creo que me arriesgaré con el pez espada...



Razonamiento Basado en Casos

Anexo: RBC como modelo del razonamiento humano

En este ejemplo se ha razonado sobre casos previos para distintos objetivos:

- ✿ Como medio para resolver un problema nuevo: sugerir pastel de tomate como plato principal.
- ✿ Para sugerir una adaptación para una solución imperfecta: sustituir el queso por tofu.
- ✿ Para advertir sobre posibles fallos: en una ocasión Elena no quiso comer pescado.
- ✿ Para interpretar una situación: ¿Por qué Elena come pescado en unas ocasiones y en otras no?



Razonamiento Basado en Casos

Anexo: RBC como modelo del razonamiento humano

Los seres humanos resolvemos muchos problemas razonando con casos previos:

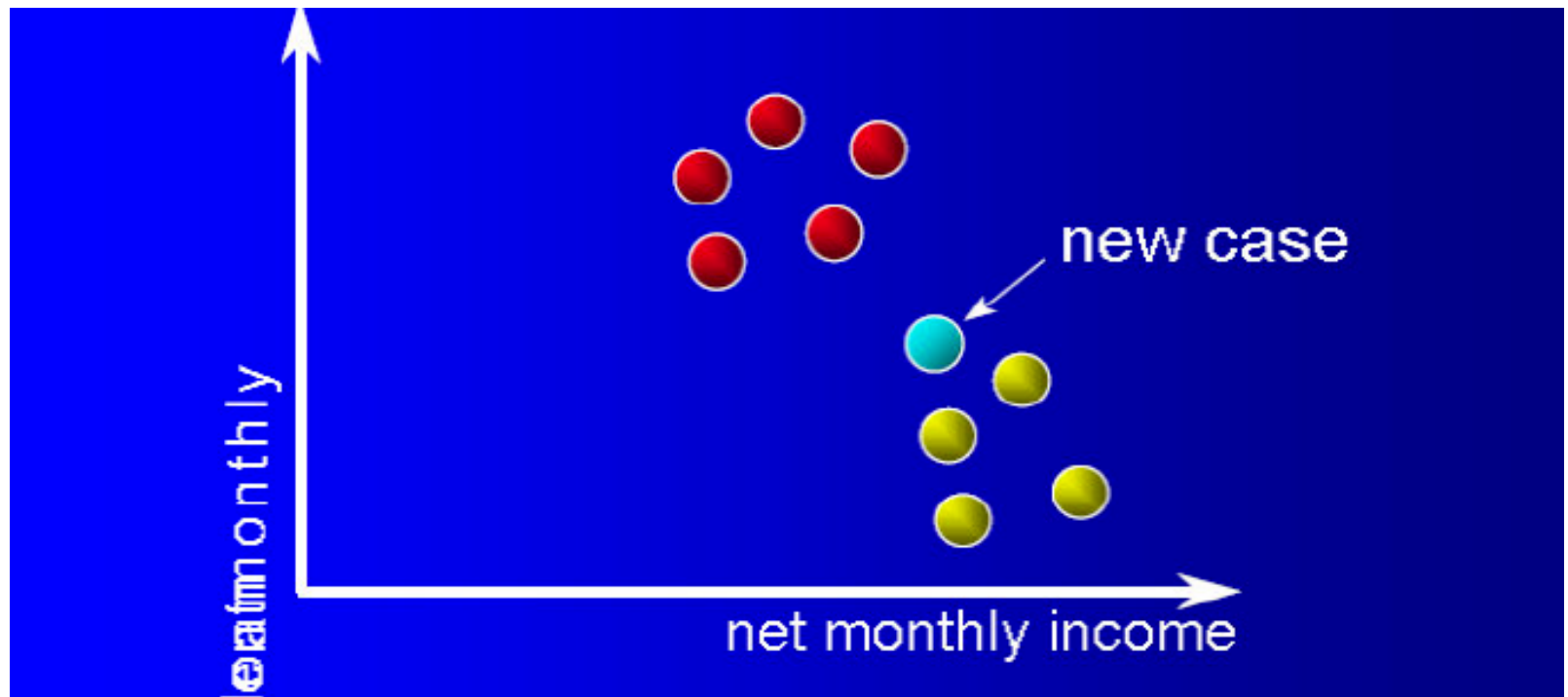
- ✿ Los abogados usan en sus argumentaciones el veredicto de casos previos – sobre el sistema judicial anglosajón-
- ✿ Los médicos buscan conjuntos de síntomas que identifican al paciente con algún conjunto de casos previos
- ✿ Los ingenieros toman muchas de sus ideas de soluciones previas ya construidas con éxito
- ✿ Los programadores expertos reutilizan esquemas más o menos abstractos de las soluciones
- ✿ En la vida diaria, donde la mayor parte de las tareas son de una u otra forma repetitivas: ir a un establecimiento, confeccionar el menú de la semana, organizar un viaje....

Plausibilidad Psicológica: El RBC es un modo natural de razonamiento de los seres humanos!



Razonamiento Basado en Casos

Gráfico:



Watson, Ian. Case-Based Reasoning Gentle Intro



Razonamiento Basado en Casos

Características

El RBC es:

- ✿ Intuitivo
- ✿ Relativamente simple de poner en práctica
- ✿ Transparente
- ✿ Aprende



Razonamiento Basado en Casos

Características

- ✿ En la vida real el espacio del problema es N dimensional
- ✿ Vectores de características se pueden ponderar para reflejar su importancia relativa
- ✿ Nuevas características se pueden agregar si adquieren relevancia
- ✿ Tolerante a datos con ruido y faltantes
- ✿ Recuperación por el algoritmo del Vecino más Cercano



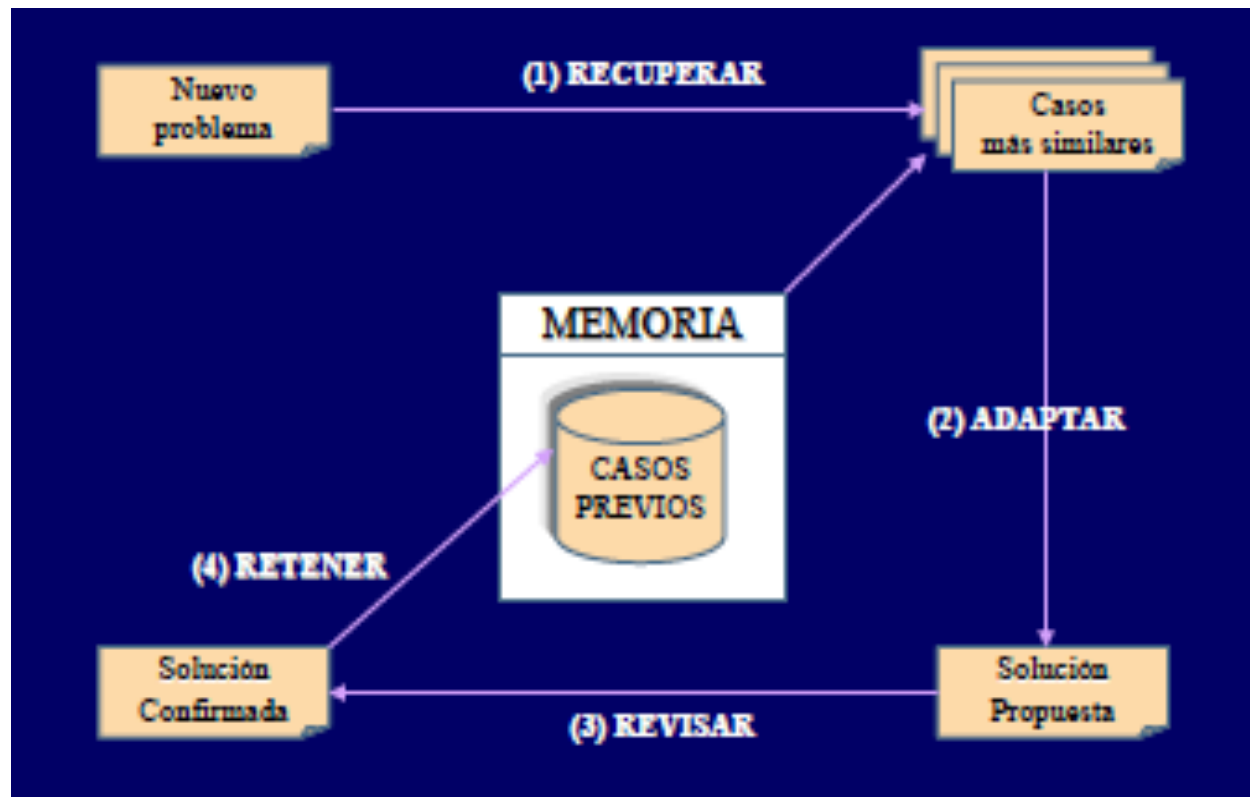
Razonamiento Basado en Casos

Tipos principales de Métodos de RBC

- ✿ Razonamiento basado en Ejemplares
- ✿ Razonamiento basado en la Memoria
- ✿ Razonamiento basado en Casos
 - ✓ Se centra en la complejidad en la representación de los casos
 - ✓ Incluye mecanismos de adaptación
 - ✓ Incluye algún tipo de conocimiento de propósito general
- ✿ Razonamiento basado en Analogías

Razonamiento Basado en Casos

Ciclo de Vida de un RBC





Razonamiento Basado en Casos

Conocimiento incluido de los Casos

¿Qué es un caso?

Un caso es una pieza contextualizada de conocimiento que representa una experiencia que enseña una lección fundamental para alcanzar los objetivos del razonador (Kolodner & Leak).

- a. El conocimiento del caso es aplicable siempre que se presente un determinado contexto.
- b. No todas las situaciones *enseñan una lección*.
 - ✿ Casos redundantes
 - ✿ Casos cubiertos por el conocimiento general
- c. ¿Qué lecciones puede enseñar un caso?
 - ✿ Cómo conseguir uno o varios objetivos
 - ✿ Cómo alcanzar el estado necesario para poder lograr un objetivo
 - ✿ Qué problemas se pueden plantear a la hora de lograr un objetivo



Razonamiento Basado en Casos

Conocimiento General

- ✿ Conocimiento de Similitud
- ✿ Reglas de Adaptación
- ✿ Generalización de Casos
- ✿ Modelos de dominio que se utilizan en las distintas fases del RBC



Análisis de Conglomerados

Medidas de Similitud

Una métrica d es una función (o regla) que asigna un número a cada par de objetos de un conjunto Ω (omega):

$$\begin{array}{ccc} \Omega \times \Omega & : & \xrightarrow{\quad d \quad} \mathbb{R} \\ (x, y) & \xrightarrow{\quad} & d(x, y) \end{array}$$



Análisis de Conglomerados

Medidas de Similitud

Lo anterior satisface las siguientes condiciones sobre los objetos x, y, z del conjunto Ω :

1. No Negatividad. $d(x, y) = 0$, si y sólo si, $x = y$
2. Simetría. Dados dos objetos x, y , la distancia d , entre ellos satisface:

$$d(x, y) = d(y, x)$$

3. Desigualdad Triangular.



Análisis de Conglomerados

Medidas de Similitud

Lo anterior satisface las siguientes condiciones sobre los objetos x, y, z del conjunto Ω :

3. Desigualdad Triangular. Para tres objetos x, y, z las distancias entre ellos satisface la expresión:

$$d(x, y) \leq d(x, z) + d(z, y)$$

La longitud de uno de los lados de un triángulo es menor o igual que la suma de las longitudes de los otros dos lados.



Análisis de Conglomerados

Medidas de Similitud

Lo anterior satisface las siguientes condiciones sobre los objetos x , y , z del conjunto Ω :

4. Identificación de no Identidad. Dados los objetos x , y :

$$\text{si } d(x, y) \neq 0, \text{ entonces } x \neq y$$

5. Identidad. Para dos elementos idénticos, x , x' , se tiene que:

$$d(x, x') = 0$$

Si los objetos son idénticos, la distancia entre ellos es cero.



Análisis de Conglomerados

Medidas de Similitud

Las medidas de similitud, de aplicación más frecuente, son las siguientes:

- ✿ Medidas de Distancia
- ✿ Coeficientes de Correlación
- ✿ Coeficientes de Asociación
- ✿ Medidas Probabilísticas de Similitud



Análisis de Conglomerados

Medidas de Similitud

- ✿ Antes de utilizar alguna de las medidas anteriores, se debe encontrar el conjunto de variables que mejor represente el conjunto de similitud, bajo el estudio a desarrollar.
- ✿ Idealmente, las variables deben escogerse dentro del marco conceptual que explícitamente se usa para la clasificación.
- ✿ La teoría en cada campo, es la base racional para la selección de las variables a usar en el estudio.



Análisis de Conglomerados

Medidas de Similitud

Las medidas de similitud, de aplicación más frecuente, son las siguientes:

- ✿ Medidas de Distancia
- ✿ Coeficientes de Correlación
- ✿ Coeficientes de Asociación
- ✿ Medidas Probabilísticas de Similitud



Análisis de Conglomerados

Medidas de Distancia

✿ Distancia Euclidiana, definida por:

$$d_{ij} = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}.$$



Análisis de Conglomerados

Medidas de Distancia

- ✿ Distancia Euclidiana
- ✿ Distancia D^2 de Mahalanobis, definida por:

$$D^2 = d_{ij} = (X_i - X_j)' \Sigma^{-1} (X_i - X_j)$$

Análisis de Conglomerados

Medidas de Distancia

- ✿ Distancia Euclidiana
- ✿ Distancia D^2 de Mahalanobis
- ✿ Distancia de Manhattan, definida por:

$$d_{ij} = \sum_{k=1}^p |X_{ik} - X_{jk}|.$$



Análisis de Conglomerados

Medidas de Distancia

- ✿ Distancia Euclidiana
- ✿ Distancia D^2 de Mahalanobis
- ✿ Distancia de Manhattan
- ✿ Distancia de Minkowsky, definida por:

$$d_{ij} = \left(\sum_{k=1}^p |X_{ik} - X_{jk}|^r \right)^{1/r} \quad \text{con } r = 1, 2, \dots$$



Análisis de Conglomerados

Medidas de Distancia

Ejemplo

Supóngase que se tienen cuatro personas cuya edad X_1 (en años), estatura X_2 (en metros), peso X_3 (en kilogramos), son los siguientes:

Persona	Edad	Estatura	Peso
A	23	1.69	61
B	40	1.70	72
C	26	1.65	68
D	38	1.68	70



Análisis de Conglomerados

Medidas de Distancia

Ejemplo

La matriz de distancias euclidianas es:

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>A</i>	0	20.25	7.62	17.49
<i>B</i>	20.25	0	14.56	2.83
<i>C</i>	7.62	14.56	0	12.17
<i>D</i>	17.49	2.83	12.17	0

Donde la distancia entre A y B, por ejemplo, resulta del siguiente cálculo:

$$d_{AB} = \sqrt{(23 - 40)^2 + (1.69 - 1.70)^2 + (61 - 72)^2} = 20.25$$

Análisis de Conglomerados

Medidas de Distancia

Ejemplo

La matriz de distancias euclidianas es:

	A	B	C	D
A	0	20.25	7.62	17.49
B	20.25	0	14.56	2.83
C	7.62	14.56	0	12.17
D	17.49	2.83	12.17	0

Se puede notar que los individuos más similares o cercanos son B y D. Resalta fácilmente de los datos.



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

- ✿ Lista de pares atributo-valor
- ✿ Bases de datos
- ✿ Sistemas de marcos y redes semánticas, modelos de memoria
- ✿ Árboles de decisión
- ✿ Árboles K-d
- ✿ Redes de activación
- ✿ Redes de recuperación de casos



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

Ejemplo

$$\text{Distancia Euclidiana}(W, T) = \sqrt{\sum_{i=1}^n (x_{iw} - x_{it})^2}$$

ID	Marca	Processor Type	Processor Speed (Mhz)	Monitor (inches)	RAM(MB)	DiscoDuro (GB)	Precio
CASE2	Compaq	Intel Celeron	1330	15	128	30	750.000
CASE1	Apple	PowerPC G4	733	13,8	128	40	1.294.000
CASE19	Apple	PowerPC G3	600	15	256	40	986.000
CASE18	Dell	Intel Pentium	1700	17	128	20	1.138.000
CASE17	Compaq	AMD Duron	700	14	256	20	1.000.000

PROBLEMAS

Marca	Processor Type	Processor Speed (Mhz)	Monitor (inches)	RAM(MB)	DiscoDuro (GB)	Precio
Compaq	Intel Celeron	1300	15	256	80	850.000
Apple	PowerPC G4	733	18	128	30	1.000.000

Encontrar el Caso más cercano, usando los valores numéricos!



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

Ejemplo

Caso 1 (níspero)	Caso 2 (naranja)	Caso 3 (manzana)	Caso 4 (albaricoque)
NARANJA ÁCIDA VERANO	NARANJA ÁCIDA INVIERNO	AMARILLA DULCE INVIERNO	NARANJA DULCE VERANO

Y si los rasgos son Categóricos?



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

Ejemplo

Caso 1 (níspero)	Caso 2 (naranja)	Caso 3 (manzana)	Caso 4 (albaricoque)
NARANJA ÁCIDA VERANO	NARANJA ÁCIDA INVIERNO	AMARILLA DULCE INVIERNO	NARANJA DULCE VERANO

Con la entrada ROJA DULCE VERANO???

Y si los rasgos son Categóricos?



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

Ejemplo

Categorías	Nro. del Caso	Elemento	Color	Sabor	Temporada
Naranja - Ácida - Verano	Caso 1	Níspero	Parcial	No	Si
Naranja - Ácida - Invierno	Caso 2	Naranja	Parcial	No	No
Amarilla - Dulce - Invierno	Caso 3	Manzana	No	Si	No
Naranja - Dulce - Verano	Caso 4	Albaricoque	Parcial	Si	Si

Con la entrada ROJA DULCE VERANO???



Razonamiento Basado en Casos

Tecnologías utilizadas en la Representación de Casos

Ejemplo

Categorías	Nro. del Caso	Elemento	Color	Sabor	Temporada
Naranja - Ácida - Verano	Caso 1	Níspero	Parcial	No	Si
Naranja - Ácida - Invierno	Caso 2	Naranja	Parcial	No	No
Amarilla - Dulce - Invierno	Caso 3	Manzana	No	Si	No
Naranja - Dulce - Verano	Caso 4	Albaricoque	Parcial	Si	Si

Con la entrada ROJA DULCE VERANO???



Universidad
Tecnológica
de Bolívar
CARTAGENA DE INDIAS



Actividad Extra-clase

Práctica



Referencias

- Atanassov, L. Antonov. Comparative Analysis Of Case Based Reasoning Software Frameworks jCOLIBRI AND myCBR. *Journal of the University of Chemical Technology and Metallurgy*, 47, 1, 2012, 83-90
- AI-CBR. www.ai-cbr.org
- Laura Lozano Javier Fernández. *Razonamiento Basado en Casos: Una Visión General*. Universidad de Valladolid España. 2006

Gracias!

