

RAONAMENT BASAT EN L'EXPERIÈNCIA

Raonament Basat en Casos

(SBC-CBR Part II – COMPONENTS DELS SISTEMES CBR)

Miquel Sànchez-Marrè

Intelligent Data Science and Artificial Intelligence Research Centre (IDEAI-UPC)

Knowledge Engineering and Machine Learning Group (KEMLG-UPC)

Computer Science Dept.

Universitat Politècnica de Catalunya · BarcelonaTech

miquel@cs.upc.edu

<http://www.cs.upc.edu/~miquel>

Course 2023/2024

<https://kemlg.upc.edu>





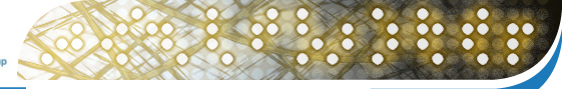
PART 2 – COMPONENTS D'UN SISTEMA CBR



Components d'un sistema CBR

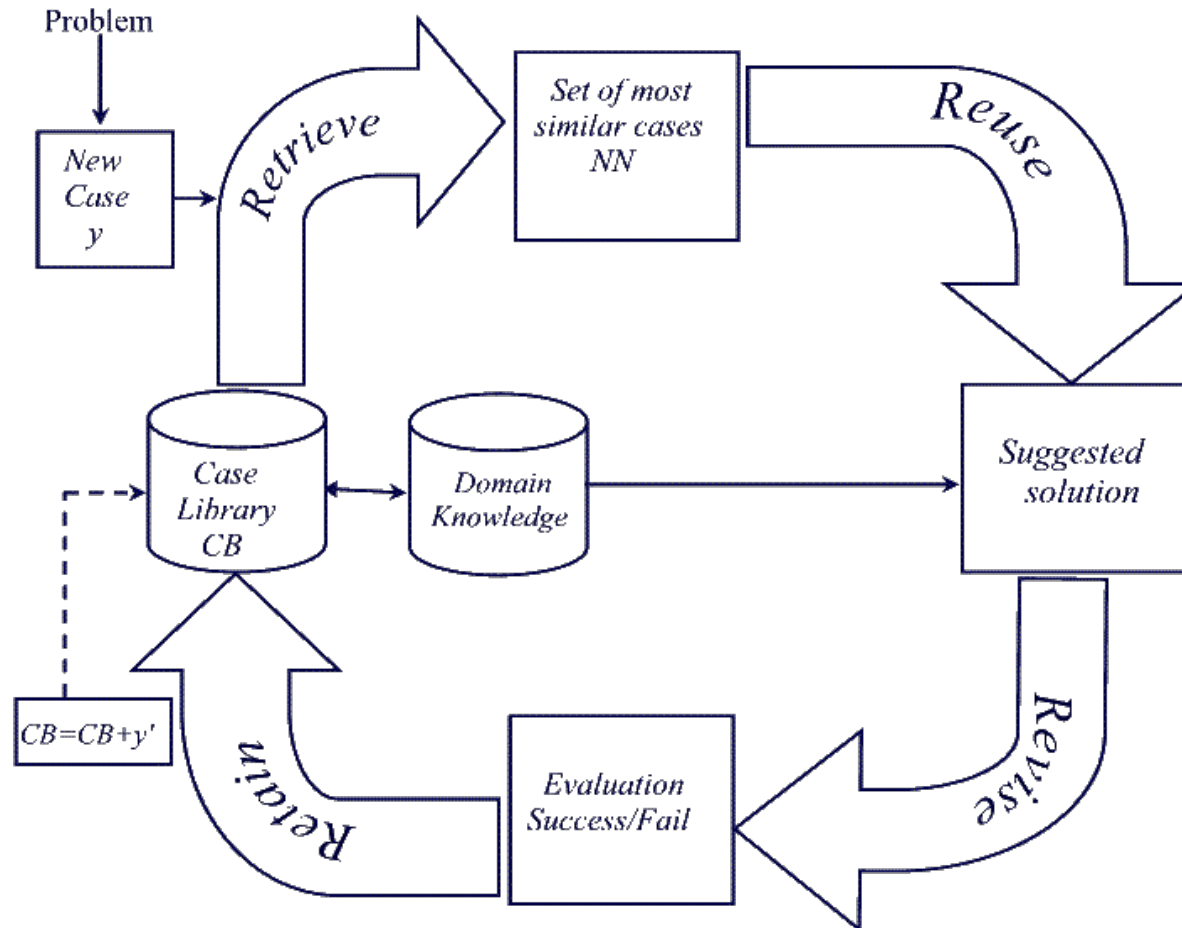
Extret de [Sánchez-Marrè, 2022]

- **Casos**
 - **Plans** o **Estructurats**
- **Biblioteca de Casos / Base de Casos / Memòria**
 - **Plana**
 - **Estructurada**
 - **Mixta**
 - **No estructurada**
- **Mètodes de recuperació**
 - **Indexació / Cerca a la biblioteca**
 - **Valoració de similaritat**
- **Mètodes d'adaptació**
- **Mètodes d'avaluació**
- **Tipus d'aprenentatge**



Les quatre R's del CBR

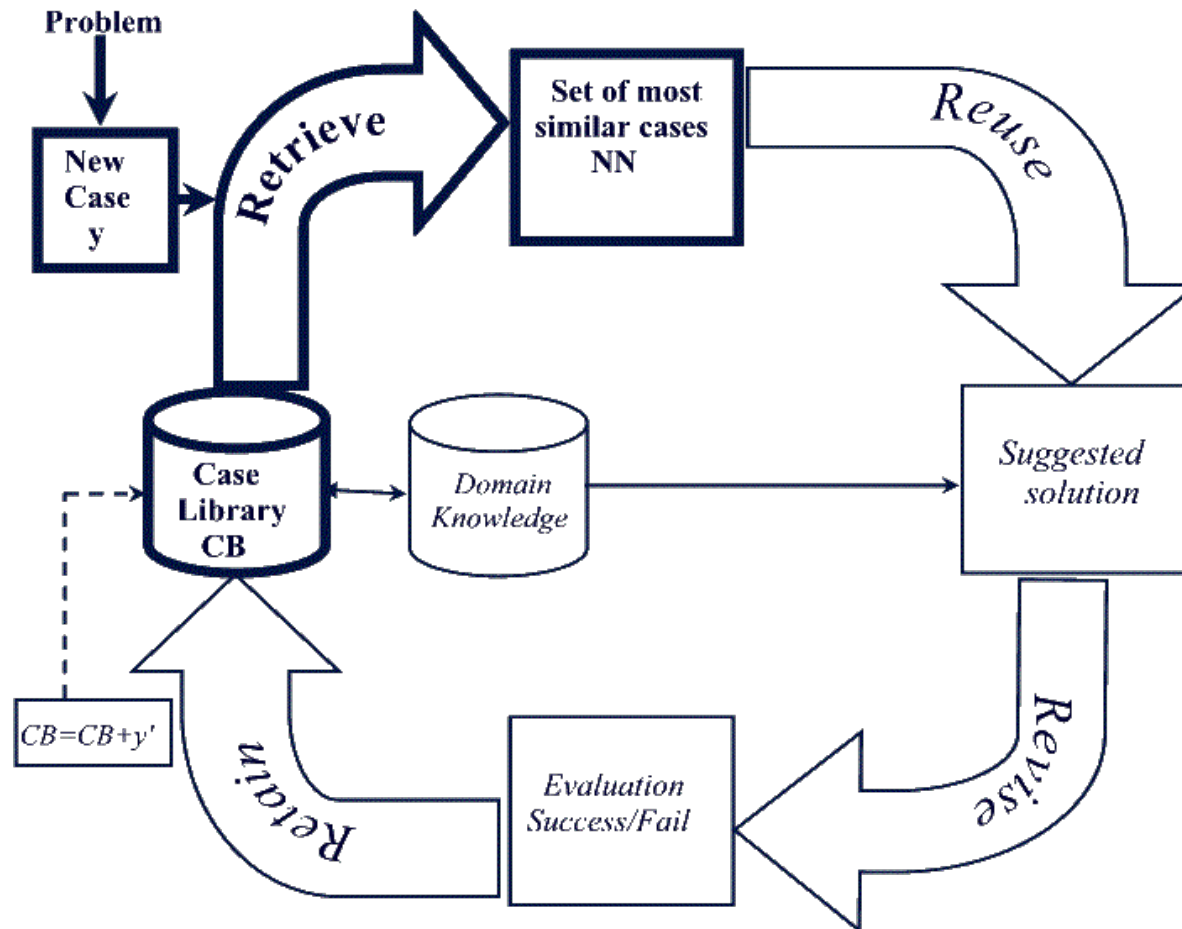
[Aamodt & Plaza, 1994]

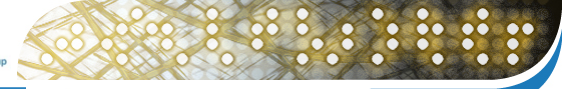




Les quatre R's del CBR

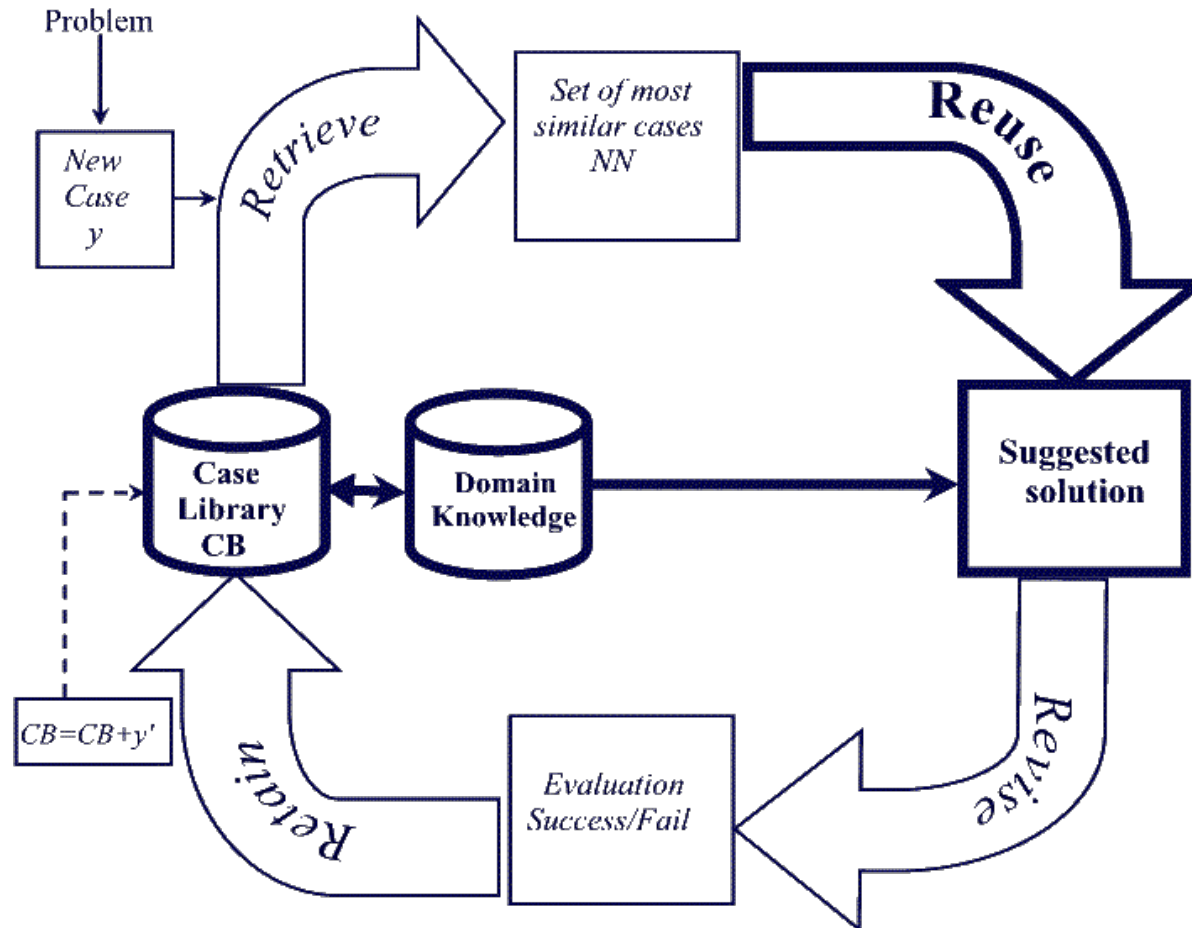
[Aamodt & Plaza, 1994]





Les quatre R's del CBR

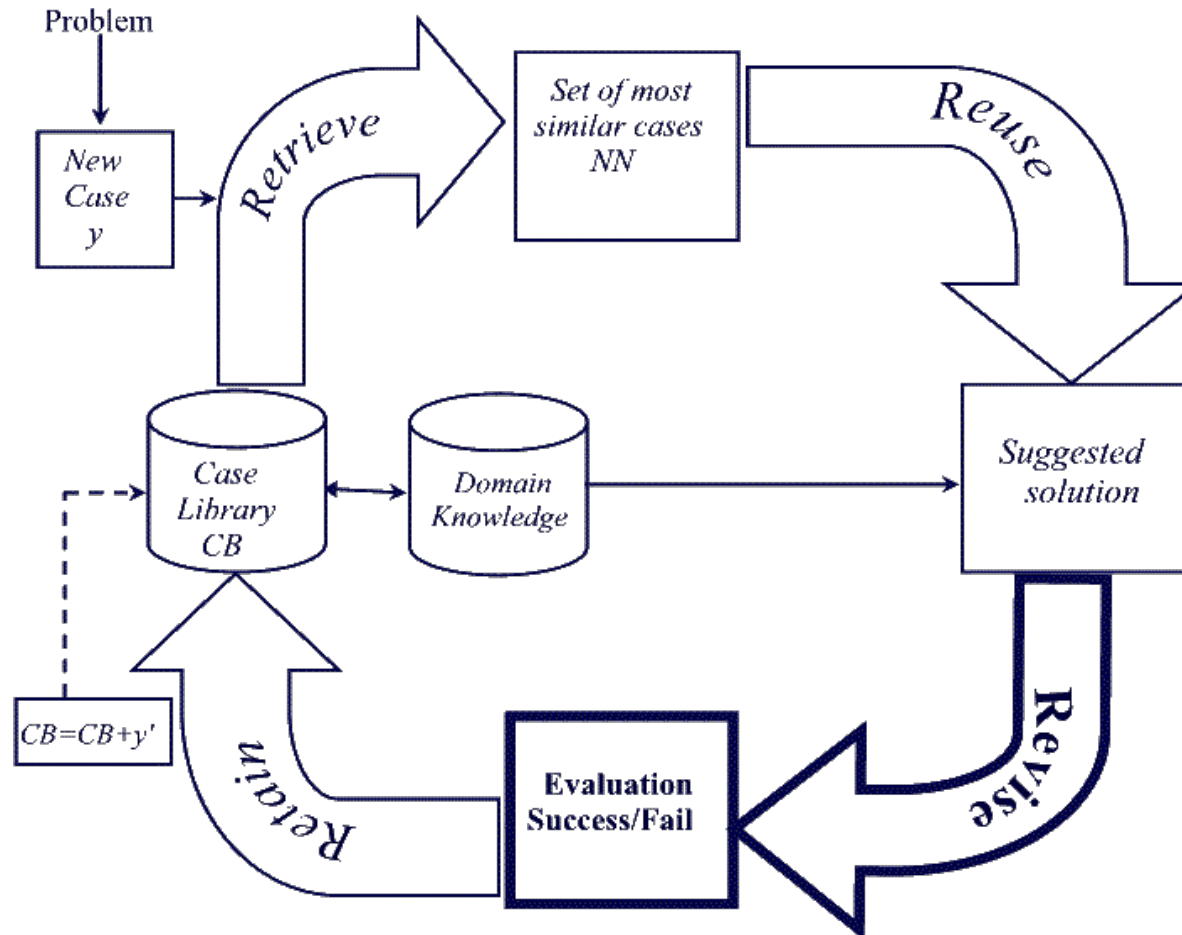
[Aamodt & Plaza, 1994]





Les quatre R's del CBR

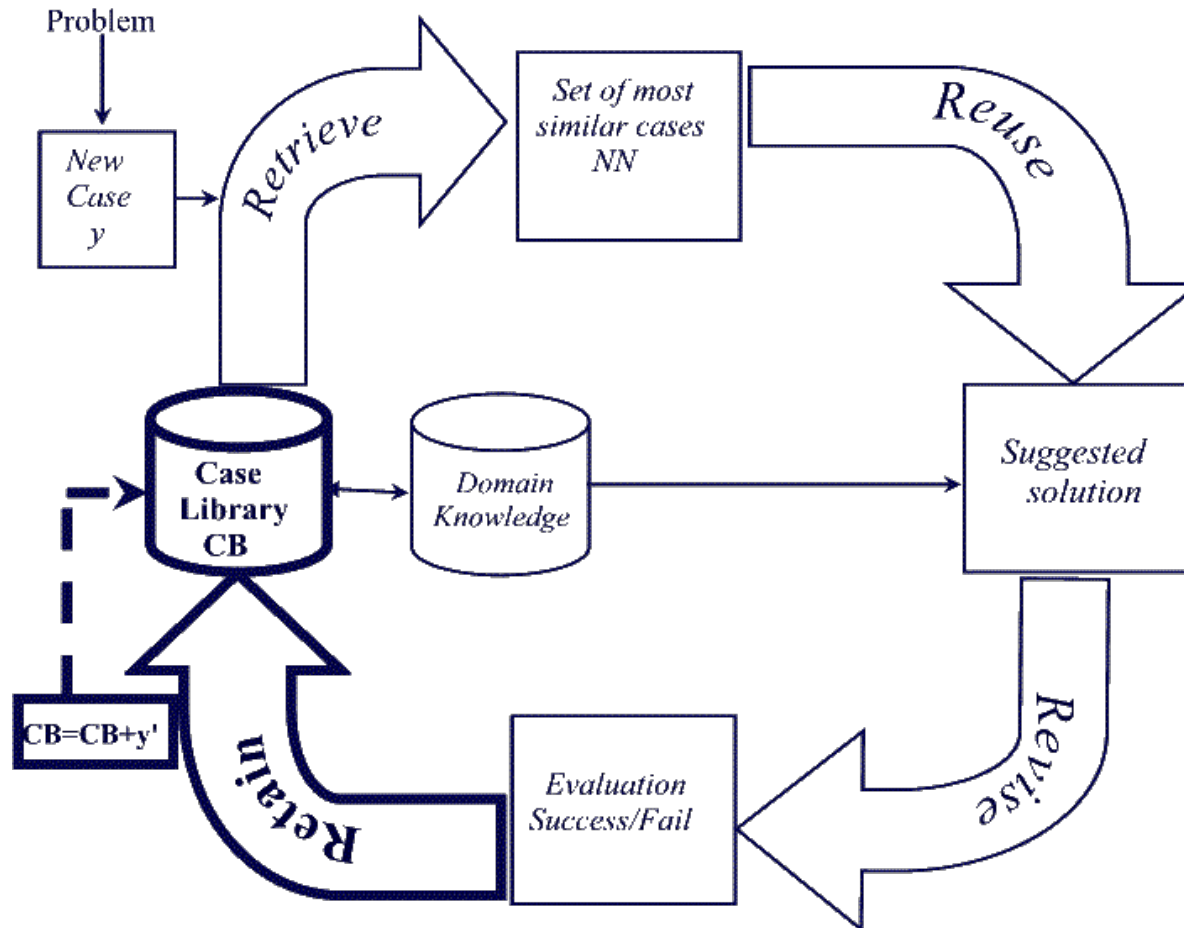
[Aamodt & Plaza, 1994]





Les quatre R's del CBR

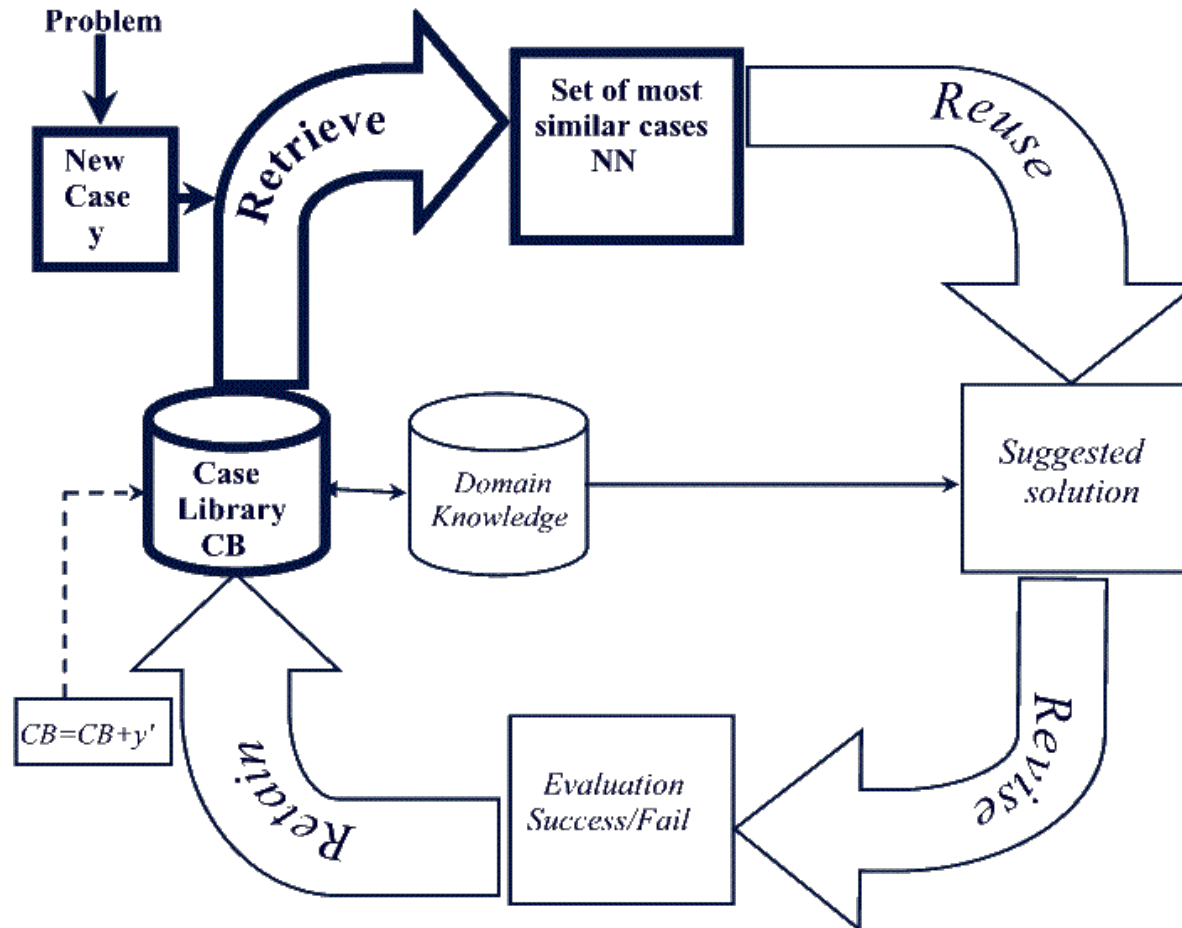
[Aamodt & Plaza, 1994]

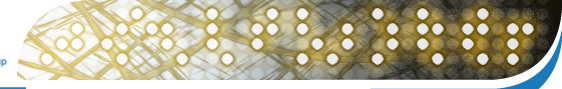




Les quatre R's del CBR

[Aamodt & Plaza, 1994]





Representació dels casos

Extret de [Sànchez-Marrè, 2022]

- **Esquemes de representació de casos més comuns**
 - **Vectors de parells atribut-valor**
 - **Estructures o Objectes**
 - **Text Lliure en Llenguatge Natural**
 - **Text i llistes de preguntes i respostes (CBR conversacional)**
 - **Grafs (arbres o xarxes)**



Representació dels casos

Vector d'atributs i valors

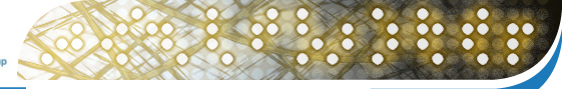
$\langle \text{attribute1}, \text{value1} \rangle$
 $\langle \text{attribute2}, \text{value2} \rangle$
 .
 .
 $\langle \text{attribute}_n, \text{value}_n \rangle$

Atributs:

$\langle \text{tipus}, \text{marca}, \text{cilindres}, \text{potencia}, \text{combustible}, \text{color}, \text{preu}, \text{any}, \text{seients} \rangle$

Casos/Exemples/Instàncies/Observacions:

$\langle \text{sport}, \text{BMW}, 8, 275, \text{gasolina}, \text{blau}, 35000, 2000, ? \rangle$



Representació dels casos

Estructures o objectes (1)

- Representació de l'estructura: un **cas** és un conjunt de característiques ("features")
 - **identificador** del cas
 - **derivació** del cas
 - **descripció** del problema (vector atribut-valor)
 - **diagnòstic/s** del problema
 - **solució** del problema
 - **evaluació** de la solució (èxit/fracàs)
 - **mesura d'utilitat**
 - altra informació relevant



Representació dels casos

Estructures o objectes (2)

- **Exemple d'un cas** en el domini de predicció volcànica i sísmica, podria ser el següent:

(:case-identifier	CASE-134
	:temporal-identifier	27/11/2004
	:case-situation-description	((SEISMIC-ACT Invaluable) (DEFORMATIONS mean-value) (GEOCHEMICAL-EVOL normal) (ELECT-PHEN level-1))
	:case-diagnostics-list	(No-eruption, Seismic-pre-Alert)
	:case-solution-plan	(Alert-Emergency-Services)
	:case-solution-evaluation	correct
)		



Representació dels Casos

Text lliure en Llenguatge Natural (1)

- En **algunes aplicacions** com:
 - Casos judicials [Weber-Lee et al., 1998]
 - Categorització de textos
 - Comerç electrònic
 - Respostes a sistemes de preguntes freqüents (FAQ's) [Lenz & Burkhard, 1997]
 - Informes mèdics i tècnics
- Cada **cas es descriu** mitjançant **frase que es busca que inclogui paraules que siguin prou bones per a discriminar el cas i que representin el problema del domini de manera fidel**
- **Exemple**
 - "A red sports car with 300 horse power and 12 cylinders, with a price less than 250000 and must be BMW"



Representació dels casos

Text Lliure en LN: respostes a FAQ's [Lenz and Burkhard, 1997] (2)

- La **base de casos** consisteix en **un conjunt de textos que contenen possibles respostes a les preguntes** fetes per l'usuari.
- El sistema descrit pren la pregunta de l'usuari (Q) expressada en llenguatge natural i recupera el text que millor coincideix com a resposta a la pregunta formulada (F).
- Per a cada cas, s'identifica un conjunt **d'Entitats d'Informació (IE)** produïdes per les paraules clau del text. Es fa una revisió de la pregunta per identificar els EI dins d'ella. La similitud s'avalua segons:

$$SIM(Q, F) = \sum_{e_i \in Q} \sum_{e_j \in F} sim(e_i, e_j)$$

on e_i i e_j son les IEs extrems de la pregunta Q i el cas F respectivament



Representació dels casos

Text Lliure en LN: respostes a FAQ's [Lenz and Burkhard, 1997] (3)

- La similaritat local **sim** compara dues IEs, i.e, dos símbols que representen un cert significat al domini. D'aquesta manera **dos casos podrien ser molt similars tot i que es representin mitjançant paraules totalment diferents però que es mapegin a IEs semblants.**
- La similitud local es defineix de la manera següent:

$$sim(e_i, e_j) = \begin{cases} 1 & \text{si } e_i = e_j \\ \sigma_1 & \text{if } e_i \text{ and } e_j \text{ represents general synonymous items} \\ \sigma_2 & \text{if } e_i \text{ and } e_j \text{ represents specific to the system synonymous items} \\ 0 & \text{otherwise.} \end{cases}$$

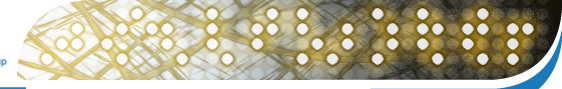
on σ_1 s'hauria d'extreure d'un diccionari, i σ_2 pot estar definida per un expert, i $0 \leq \sigma_1 \leq \sigma_2 \leq 1$.



Representació dels casos

CBR Conversacionals (1)

- Hi ha moltíssimes aplicacions de consult a la **web**.
- En aquestes aplicacions **la persona usuària introdueix un petit text** explicant el seu problema.
 - Amb aquest text el sistema **troba el cas més similar dins la seva base de casos**.
 - Llavors, presenta **un conjunt de preguntes associades al** cas que s'ha recuperat.
 - La persona usuària **pot contestar aquestes preguntes** de manera ràpida i directa.
- Aquest sistemes es coneixen com **CCBR (Conversational CBR)**. Han d'inferir automàticament els detalls que descriuen el problema a partir del text que ha introduït l'usuari.
- Al llarg de la **conversa**, el sistema avalua i mostra els casos més semblants i les seves solucions de manera progressiva fins a trobar la solució més apropiada segons criteri de l'usuari.



Representació dels casos

CBR Conversacionals (w)

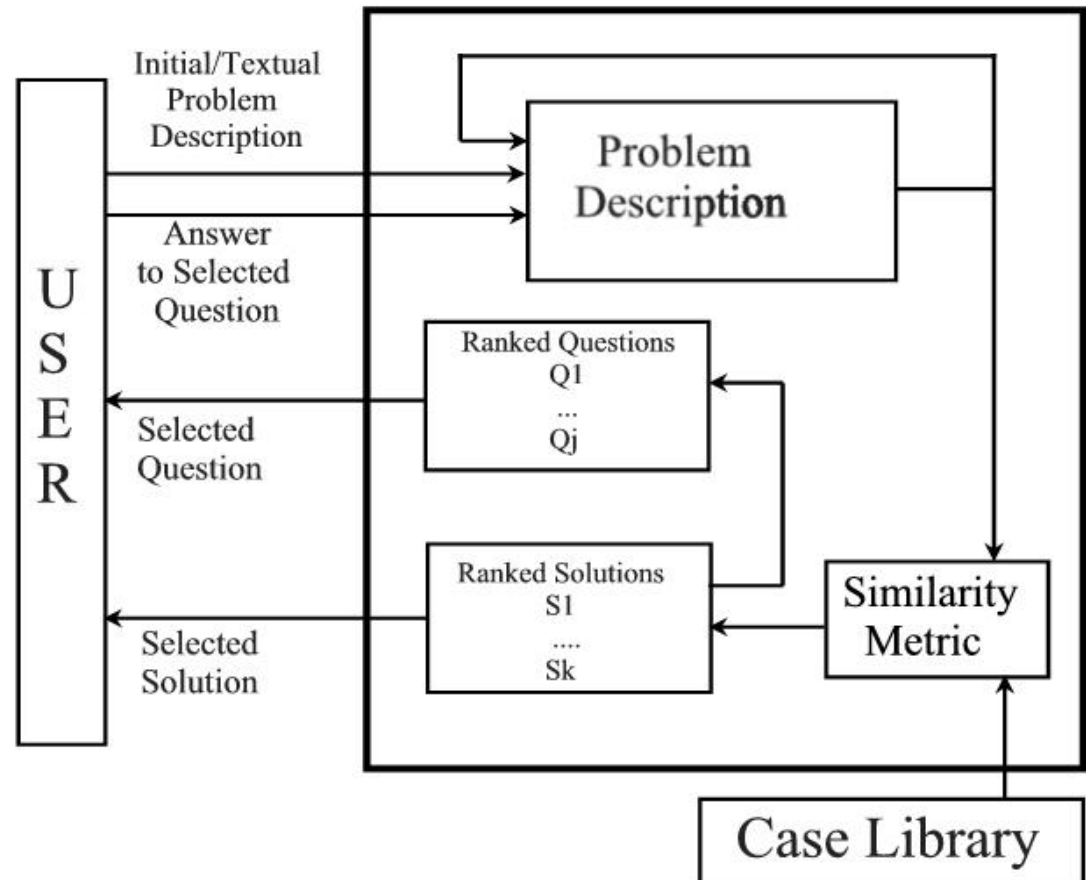
- En aquests sistemes [Aha et al., 1999], **cada cas x es representa així:**
 - Problema** $x_p = x_d + x_{qa}$: Codifica el problema resolt amb t x_s . On: la descripció x_d és un fragment de text lliure que descriu parcialment el problema d' x . L'especificació x_{qa} és un conjunt de parells <pregunta, resposta>.
 - Solució** $x_s = \{x_{a1}, x_{a2}, \dots\}$ és una seqüència d'accions x_{ai} per a contestar x_p .
- Les accions** poden ser text lliure, hyperlinks or altres objectes.
- La descripció del problema i l'especificació de cada cas actuen com el seu índex. Les preguntes en l'especificació d'un cas poden ser internament disjuntives (i.e., tenir múltiples respostes).
- El casos són exemples "positius": se suposa que aplicant x_s a x_p tindrà èxit. x actua com a un prototipus per resoldre consultes (preguntes) l'especificació del problema de les quals és similar a la d' x .

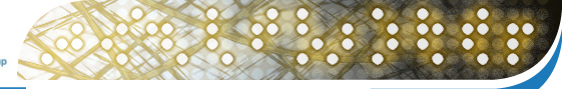


Case Representation

Conversational CBR (3)

Solució de
problemes
amb CCBR

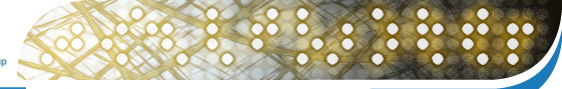




Representació dels casos

Grafs (arbres / xarxes) (1)

- En alguns dominis, o bé els **objectes o les característiques que representen el problema estan altament relacionades les unes amb les altres. Però:**
 - Aquesta relació no es pot representar de manera eficient en cap dels models descrits anteriorment.
 - En aquests dominis, es podrien utilitzar els **grafs** com a formalisme de representació de casos.
- (*com ja sabeu...*) un **graf** es defineix com una estructura $G = \langle VG, AG \rangle$
 - VG és un conjunt finit de vèrtexs
 - $AG \subseteq VG \times VG$ és un conjunt d'arestes.
 - Els vèrtexs s'utilitzen per representar els objectes del domini
 - Les vores expressen les relacions i les restriccions entre elles
- **Exemples:**
 - La planificació de l'horari d'un curs escolar [Burke et al., 2001].
 - En alguns dominis, les arestes poden representar predicats binaris com en CHIRON i CAPER [Sanders et al., 1997].

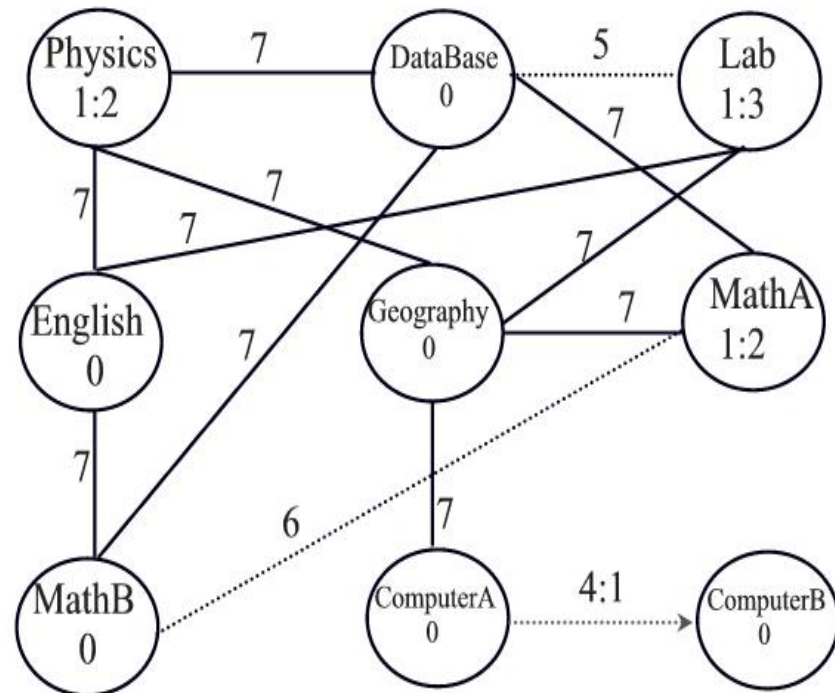


Representació dels casos

Grafs (arbres / xarxes) (2)

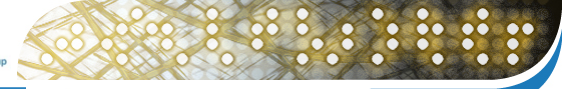
Un **problema de creació d'horaris**
representat amb un graf d'atributs

hard constraint _____
soft constraint



arestes etiquetades 7 \equiv no simultani
arestes etiquetades 6 \equiv no consecutiu
arestes etiquetades 5 \equiv consecutiu
arestes etiquetades 4 \equiv abans de

x:y significa \equiv x és l'etiqueta i y representa el valor de l'atribut
0 significa un curs normal
1:n significa un curs múltiple, n vegades per setmana



Organització de la base de casos

Extracted from [\[Sánchez-Marrè, 2022\]](#)

- **Memòria plana**
- **Organitzacions estructurades**
 - **Organitzacions jeràrquiques o indexades**
 - ♦ Arbres/Grafs de característiques compartides
 - ♦ Arbres/Grafs Discriminants
 - ♦ Arbres/Grafs Redundants Discriminants
 - ♦ k-d trees
 - Organitzacions Orientades a Objectes
 - Bases de casos múltiples
- **Organitzacions mixtes**
 - Esquemes d'indexació + estructures planes
- Organitzacions no estructurades
 - Text no estructurat

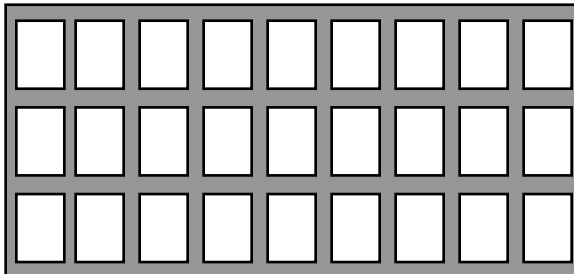


Organització de les bases de casos

Representacions habituals

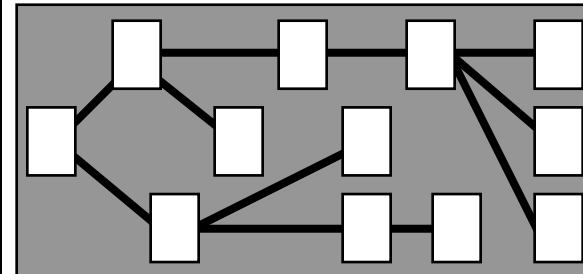
- **Memòria plana**
- **Organitzacions Estructurades**
 - **Organitzacions Jeràrquiques/Indexades**

Flat memories



- Easy to manage
- Slow for retrieval
- Always finds the best

Hierarchical memories



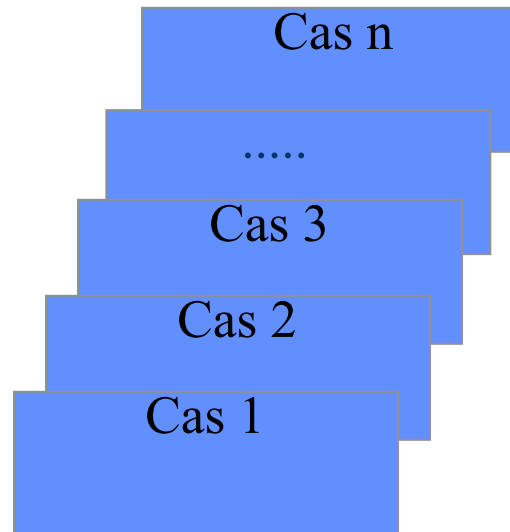
- Hard to manage
- Fast for retrieval
- Heuristic search

- **Organitzacions Mixtes**
 - **Esquemes d'indexació + estructures planes**



Organització de les Base de Casos

Memòria plana (1)



Quan s'utilitza una **memòria plana** com a estructura de base de casos i vectors de parells atribut-valor com a estructura de representació, llavors el CBR s'anomena habitualment com:

Sistemes IBL (**Instance-Based Learning**)

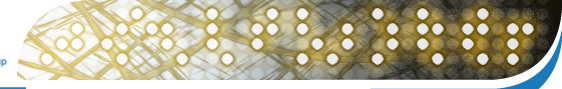


Organització de les Base de Casos

Memòria plana (2)

Exemple:

esportiu, BMW, 8, 275, gasolina, blau, 35000, 2000, ?
esportiu, AstonM, 12, 450, gasolina, vermell, 120000, 2002, 2
esportiu, Maserati, 12, 380, gasolina, verd, 80000, 2002, 2
familiar, Citroën, 4, 95, diesel, vermell, 20000, 2002, 5
familiar, Renault, 4, 65, diesel, blanc, 12000, 2002, 5



Organització de les Base de Casos

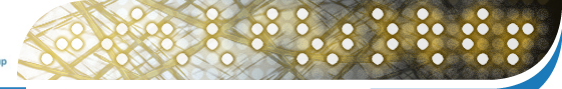
Memòria plana (3)

Algoritme de cerca en memòria plana

Entrada S: Conjunt de casos de la base de casos, y: Nou cas
Sortida r: index del cas més proper

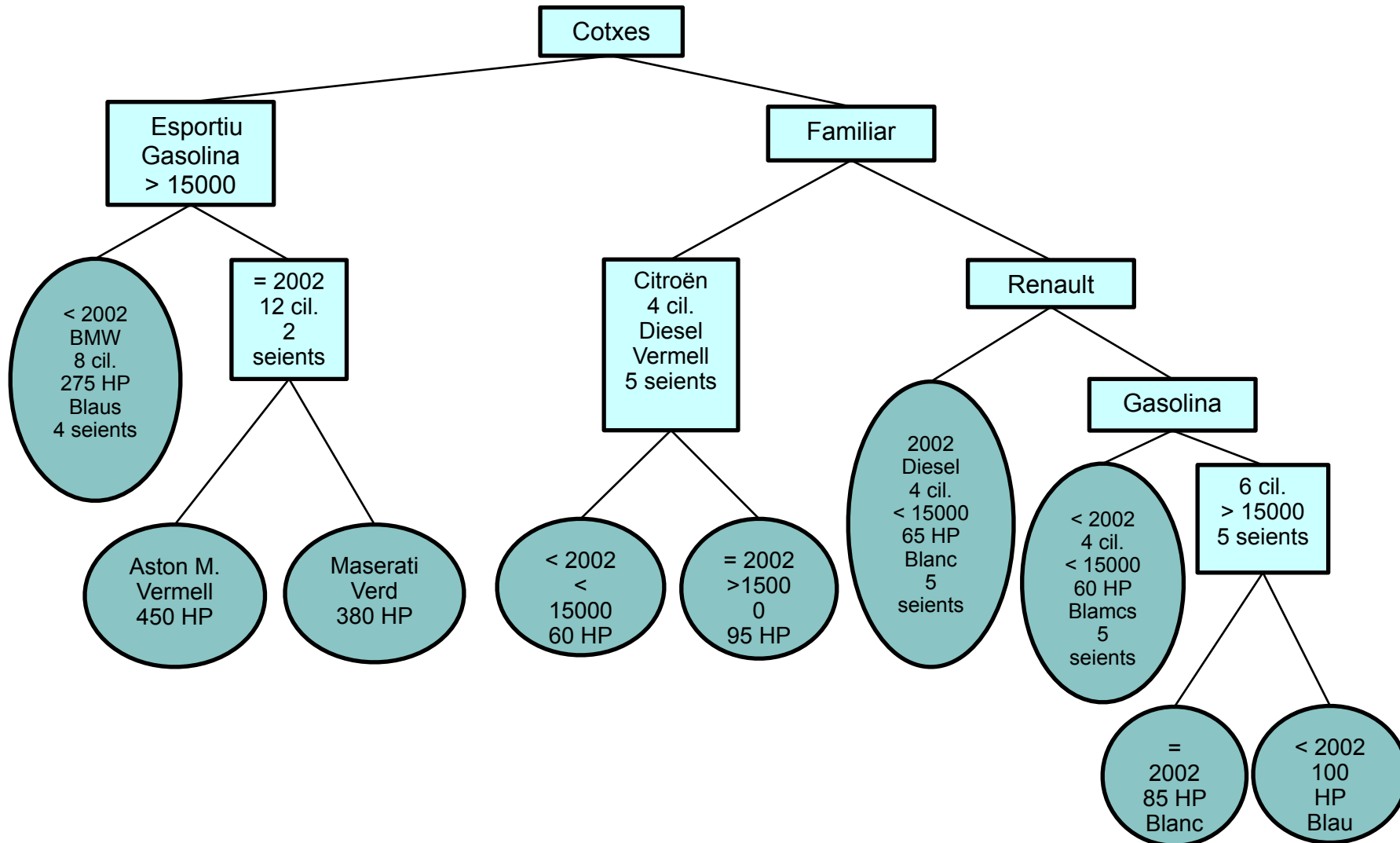
```

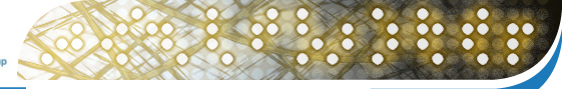
N ← |S|
millor_disimilaritat ← +∞
i ← 0
repetir
{
    i ← i+1
    dissimilaritat_actual ← dissimilaritat(Si, y)
    If dissimilaritat_actual ≤ millor_disimilaritat I
        r ← index(Si)
        millor_disimilaritat ← dissimilaritat_actual
    endIf
}
fins que i=N
  
```



Organització de les bases de casos

Arbres/Xarxes de característiques compartides (1)





Organització de les bases de casos

Arbres/Xarxes de característiques compartides (2)

Algoritme per a construir una xarxa de característiques compartides

Escollir un mètode de clustering

1. Crear el node arrel de l'arbre, N .
2. Sigui C un conjunt de casos que cal organitzar
3. Posar els valors de qualsevol característica que comparteixen **tots** els casos de C a N
4. Partir C segons el mètode de clustering. Crear un node per a cada partició convertint aquests nodes en successors d' N
5. Per a cada partició:
 - (a) Crear un node N_i
 - (b) Si N_i conté més d'un cas, repetir el pas 4, amb $N=N_i$, C =els casos de la partició.
 - (c) Sinó, posar les característiques de l'únic cas al node N_i .



Organització de les bases de casos

Arbres/Xarxes de característiques compartides (3)

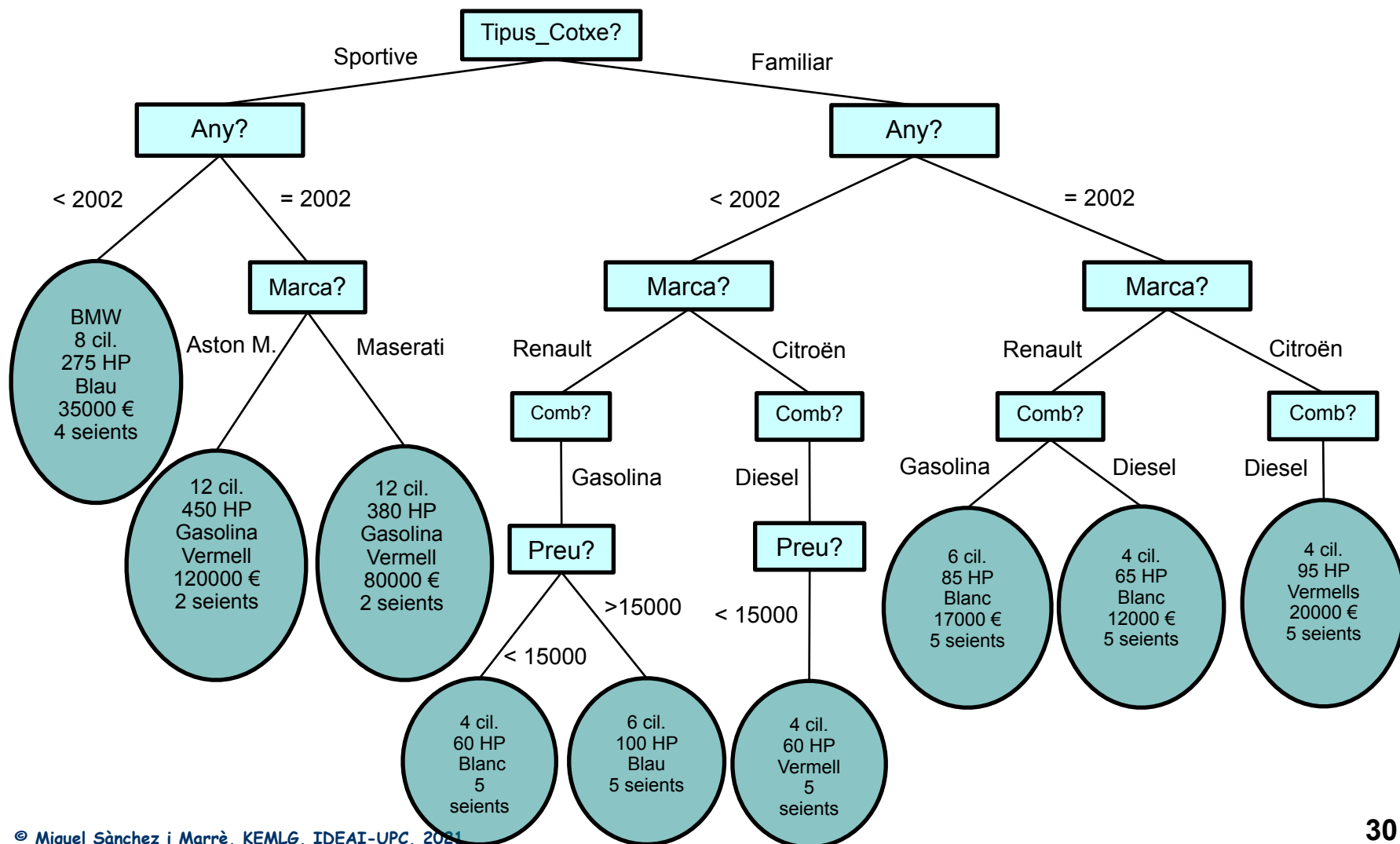
Algoritme de cerca en una xarxa de característiques compartides

1. Sigui N = arrel
2. Repetir fins que N sigui un cas:
3. Trobar el node sota N que millor concorda ("match") amb l'entrada.
4. Retornar N



Organització de les bases de casos

Arbres/Xarxes de característiques discriminants (1)



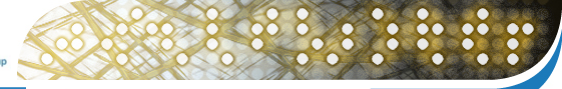


Organització de les bases de casos

Arbres/Xarxes de característiques discriminants (2)

Algoritme per construir un arbre/xarxa discriminant

1. Sigui N = arrel de l'arbre.
Sigui C = un nou cas a afegir.
Sigui P = llistes de preguntes prioritzades per als nodes
2. Si el node N té un cas, guardar com $C1$, fer que N preguntí la següent pregunta de P .
Per a cada resposta A a la pregunta, afegir un subnode a N amb A amb la seva pregunta i C o $C1$ com el seu valor (a menys que C i $C1$ donin la mateixa resposta A ; llavors N = nou node, N conté $C1$, i repetir aquest pas).
3. Sinó (si N no es un cas):
Pregunta la pregunta al node N de C . Sigui A = la seva resposta.
Si ja existeix un subnode amb A com a resposta fer que aquest node= N , i anar al pas 2.
4. Sinó (si no hi ha subnode amb A com a resposta):afegiu un subnode a N amb A com a resposta a la pregunta a N i poseu C en aquest node.



Organització de les bases de casos

Arbres/Xarxes de característiques discriminants (3)

Algoritme per buscar en una xarxa discriminant

1. Sigui N = arrel de l'arbre.
2. Repetir fins que N sigui un as:
 - a. Preguntar la pregunta N de l'entrada
 - b. Si N = subnode amb la resposta que millor coincideixi amb l'entrada.
3. Retorna N .



Organització de la Base de Casos

Arbres/Xarxes Discriminants Redundants

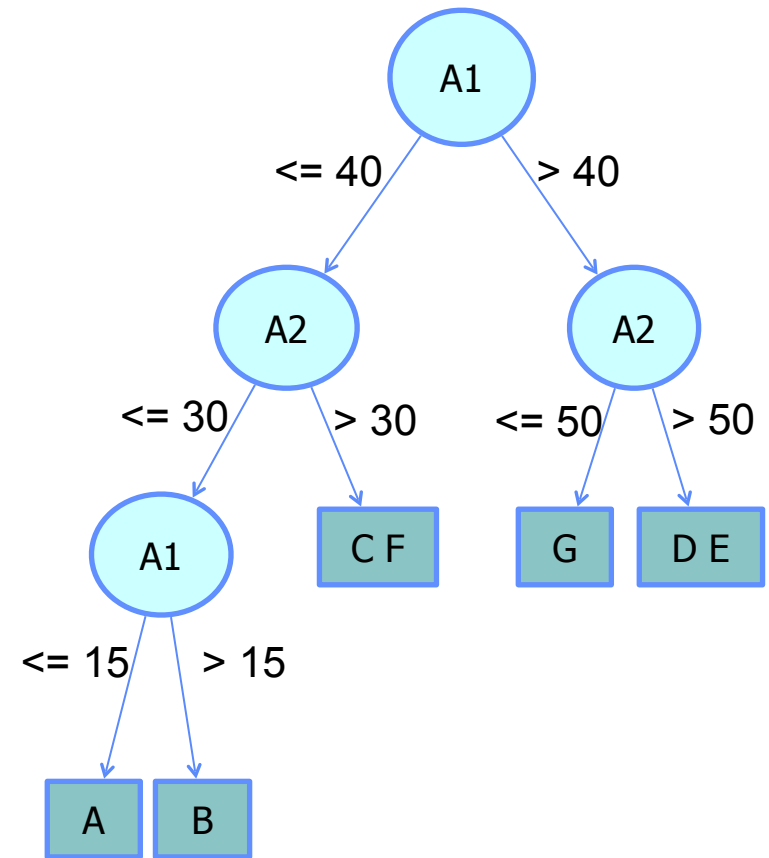
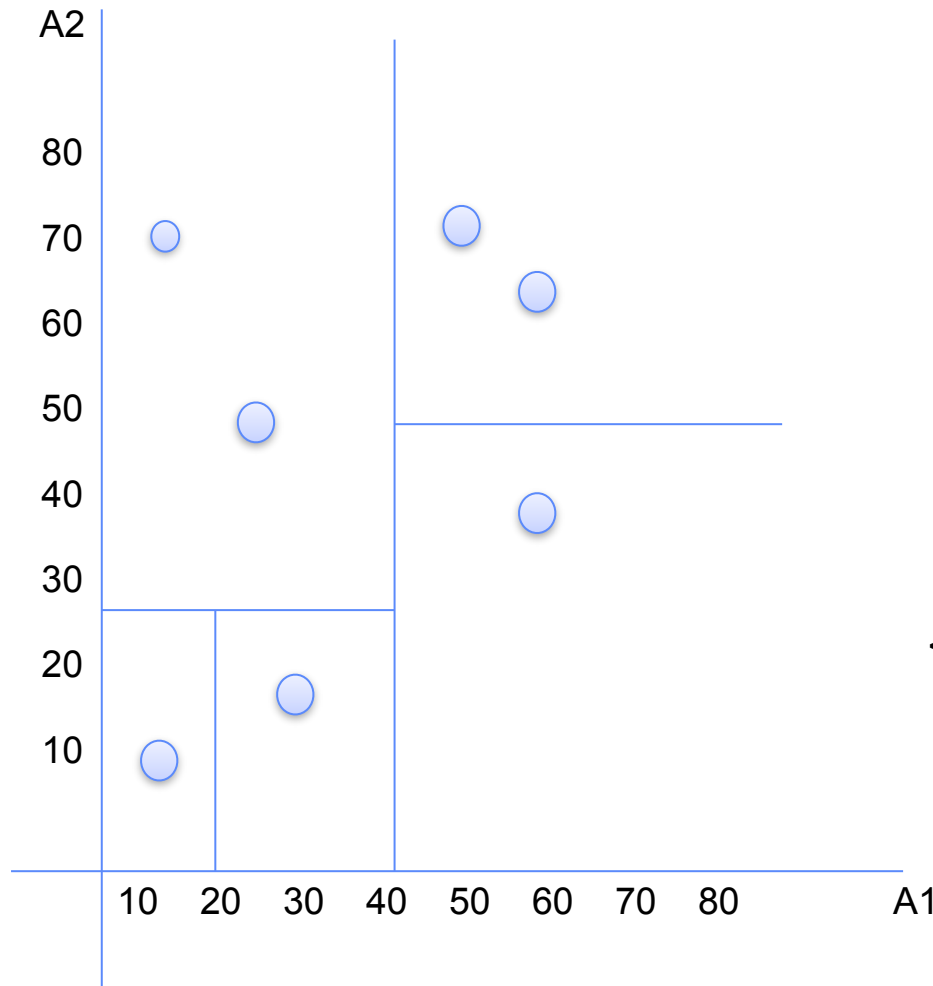
Xarxa discriminant redundant

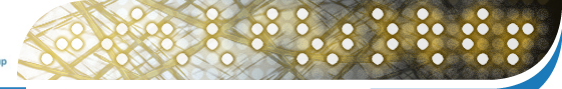
- Un **conjunt de xarxes discriminants** en comptes d'una sola i cada xarxa amb un **ordre diferent** de preguntes.
- La cerca és pot fer en **paral·lel**. Si en una xarxa la resposta a una pregunta no hi és, s'acaba cerca en aquesta xarxa i es continua dins les altres.
- Degut a la redundància, almenys en una de les xarxes, normalment, es trobarà un cas concordant, si és que n'existeix algun.
- Per mantenir la redundància sota control aquesta xarxes **es combinen normalment amb xarxes de característiques compartides**.



Organització de les bases de casos

Arbres k-d (k-d trees)





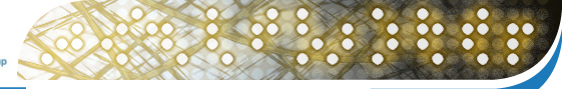
Anàlisi del diversos formalismes

Memòria Plana

- **Avantatges**
 - Implementació fàcil
 - Tasques d'aprenentatge i oblit fàcils
 - Sempre recupera el millor cas ("cas òptim")
- **Inconvenients**
 - Alt cost computacional quan augmenta la base de casos
 - Mala estructuració i interpretació de la Base de Casos

Estructures jeràrquiques

- **Avantatges**
 - Cerca ràpida
 - Estructuració i interpretació major de la Base de Casos
- **Inconvenients**
 - No sempre recupera el millor cas ("cas òptim")
 - Necessitats de major capacitat d'emmagatzematge
 - Aprendre i oblidar és una mica més complex



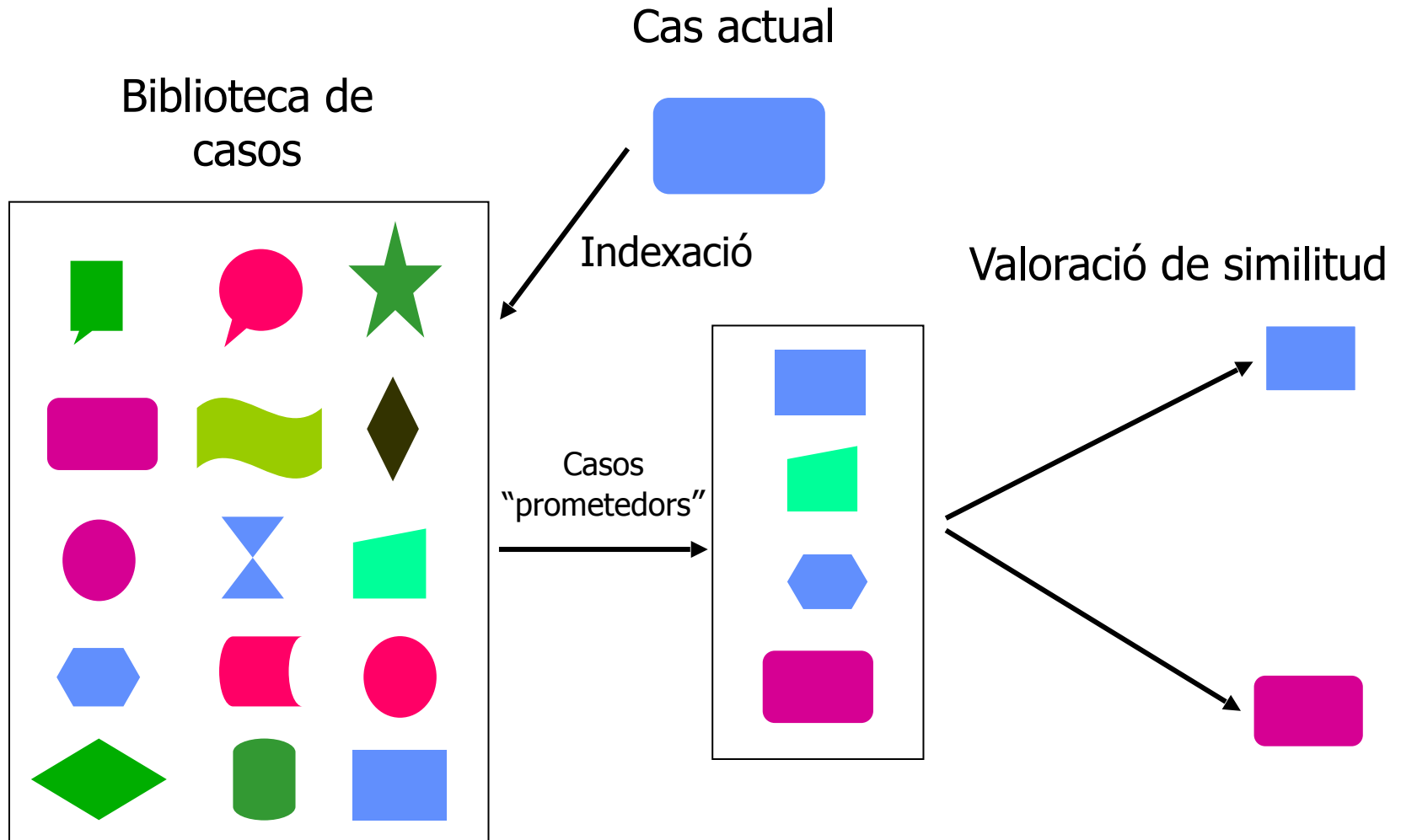
Recuperació ("Retrieval")

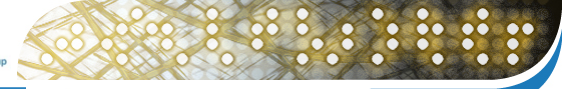
Extret de [Sánchez-Marrè, 2022]

- **La recuperació de casos és més difícil que** en Bases de Dades
 - Recuperació a BD = concordància exacta
 - Recuperació a CBR = concordància parcial (**grau de similaritat**)
- El rendiment del procés de recuperació depèn en gran mesura del **Organització de la Base de Casos**
- **L'estructura de la biblioteca de casos i la representació dels casos** faciliten la recuperació de casos rellevants i compararlos amb el problema actual
- **Passos de recuperació**
 - **Indexació**: cercar dins l'estructura de la biblioteca de casos un conjunt de "casos prometedors"
 - **Valoració de Similitud**: determinar el valor exacte de similitud per a cadascun dels casos cercats



Recuperació





Aspectes importatns que afecten el pas de recuperació

- **Cerca**
 - **Formalisme de Representació dels Casos**
 - ♦ **Vectors Atribut-Valor**
 - ♦ **Estructures**
 - ♦ **Text lliure de Llenguatge Natural**
 - ♦ **Grafs**
 - ♦ **etc.**
 - **Indexació de la Llibreria de Casos**
 - ♦ **Plana**
 - ♦ **Jeràrquica**
- **Valoració de la similitud**
 - **Mesures de semblança**
 - **Discretització d'atributs continus**
 - **Valors que falten**
 - **Rellevància dels atributs**



Mesures de similitud (1)

- Derivades de la mètrica de Minkowski**

$$d(x, y) = \left(\sum_{k=1}^K |x_k - y_k|^r \right)^{1/r} \quad r \geq 1$$

- Manhattan**

$$d(x, y) = \frac{\sum_{k=1}^K w_k * d(x_k, y_k)}{\sum_{k=1}^K w_k}$$

$$d(x_k, y_k) = \begin{cases} |x_k - y_k| & \text{if } k \text{ is continuous} \\ 0 & \text{if } x_k = y_k \text{ and } k \text{ is discrete} \\ 1 & \text{if } x_k \neq y_k \text{ and } k \text{ is discrete} \end{cases}$$

- Euclídea**

$$d(x, y) = \sqrt{\frac{\sum_{k=1}^K w_k^2 * (atr_dist(x_k, y_k))^2}{\sum_{k=1}^K w_k^2}}$$

$$atr_dist(x_k, y_k) = \begin{cases} |x_k - y_k| & \text{if } k \text{ is continuous} \\ 0 & \text{if } x_k = y_k \text{ and } k \text{ is discrete} \\ 1 & \text{if } x_k \neq y_k \text{ and } k \text{ is discrete} \end{cases}$$

- Mesures Normalitzades**

- Clark** [Lance and Williams, 1966]

$$d(x, y) = \sum_{k=1}^K \frac{|x_k - y_k|^2}{|x_k + y_k|^2}$$

- Canberra** [Lance and Williams, 1966]

$$d(x, y) = \sum_{k=1}^K \frac{|x_k - y_k|}{|x_k + y_k|}$$



Mesures de similitud (2)

- Mesures probabilístiques**

- Value Difference Metric (VDM)** [Stanfill and Waltz, 1986]

$$d(x, y) = \sum_{k=1}^K w_k(x_k) * \delta(k, x_k, y_k) \quad w_k(x_k) = \sqrt{\sum_{i=1}^c \left(\frac{N(x_k, c_i)}{N(x_k)} \right)^2} \quad \delta(k, x_k, y_k) = \sum_{i=1}^c \left(\frac{N(x_k, c_i)}{N(x_k)} - \frac{N(y_k, c_i)}{N(y_k)} \right)^2$$

- Heterogeneous Valued Difference Metric (HVDM)** [Wilson And Martinez, 1997]

$$HVDM(x, y) = \sqrt{\sum_{k=1}^K d_k^2(x_k, y_k)} \quad d_k^2(x_k, y_k) = \begin{cases} 1, & \text{if } x_k \text{ or } y_k \text{ is missing,} \\ \text{otherwise} \\ \text{normalized_vdm}_k(x_k, y_k), & \text{if } k \text{ is discrete} \\ \text{normalized_diff}_k(x_k, y_k), & \text{if } k \text{ is continuous} \end{cases}$$

$$\text{normalized_vdm}_k(x_k, y_k) = \sqrt{\sum_{i=1}^c \left| \frac{N(x_k, c_i)}{N(x_k)} - \frac{N(y_k, c_i)}{N(y_k)} \right|^2} \quad \text{normalized_diff}_k(x_k, y_k) = \frac{|x_k - y_k|}{4\sigma_k}$$

- Bayesian Similarity Measure** [Kontkanen et al., 2000]

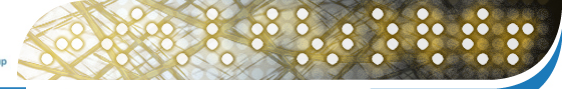
$$S(x|y) = \sum_{i=1}^c P(y, x_c = c_i | x, T, \Psi) = \sum_{i=1}^c \frac{N(c_i) + \mu_i}{N + \sum_{j=1}^c \mu_j} \prod_{k=1}^K \frac{N(x_k, c_i) + \sigma_{i,k,y_k}}{N(c_i) + \sum_{l=1}^c \sigma_{i,k,y_l}}$$

- Short & Fukunaga (SF)** [Short and Fukunaga 1981; Myles and Hand, 1990]

$$d(x, y) = \sum_{i=1}^c |p(c_i|x) - p(c_i|y)| \quad d(x, y) = \sum_{i=1}^c p(c_i|x) * |p(c_i|x) - p(c_i|y)|$$

- Minimum Risk Metric (MRM)** [Blanzieri and Ricci, 1999]

$$d(x, y) = \sum_{i=1}^c p(c_i|x)(1 - p(c_i|y))$$



Mesures de similitud (3)

- Mesures heterogènies sensibles als pesos**

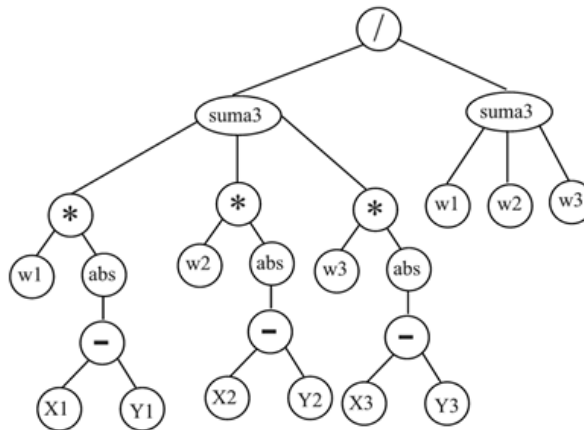
- L'Eixample** [Sánchez-Marrè et al., 1998; Sánchez-Marrè, 1996]

$$d(x, y) = \frac{\sum_{k=1}^K e^{w_k} \times d(x_k, y_k)}{\sum_{k=1}^K e^{w_k}}$$

$$d(x_k, y_k) = \begin{cases} \frac{|qtv(x_k) - qtv(y_k)|}{upperval(k) - lowerval(k)} & \text{if } k \text{ is continuous and } w_k \leq \alpha \\ \frac{|qlv(x_k) - qlv(y_k)|}{V_k - 1} & \text{if } k \text{ is continuous and } w_k > \alpha \\ & \text{or } k \text{ is ordered discrete} \\ 1 - \delta_{qlv(x_k), qlv(y_k)} & \text{if } k \text{ is non ordered discrete} \end{cases}$$

$$\delta_{qlv(x_k), qlv(y_k)} = \begin{cases} 1 & \text{if } qlv(x_k) = qlv(y_k) \\ 0 & \text{if } qlv(x_k) \neq qlv(y_k) \end{cases}$$

- Mesures Generades per Programació Genètica** [Golobardes et al., 2000]





Mesures de similitud (4)

- **Mesures de correlació**
 - **Cosine Similarity Metric** [Gordon, 1999]

$$d(x, y) = \frac{\sum_{k=1}^K x_k y_k}{\sqrt{\sum_{k=1}^K x_k^2 \sum_{k=1}^K y_k^2}}$$

- **Mahalanobis Similarity Metric**

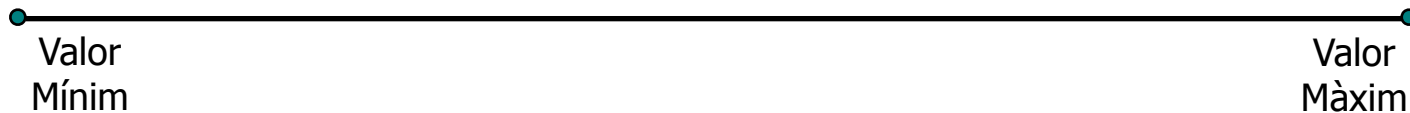
$$d(x, y) = \sqrt{(x - y)^T \Sigma^{-1} (x - y)}$$



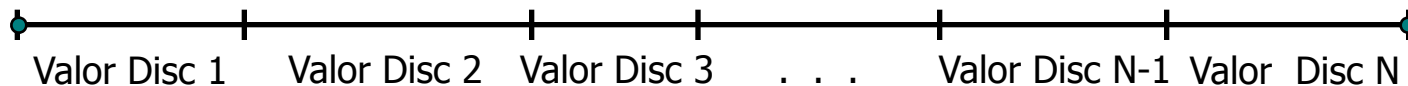
Discretització d'atributs

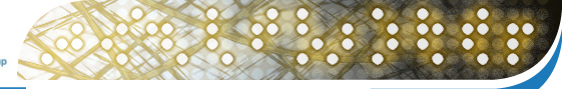
- Procés en el qual a partir dels valors continus d'un atribut, s'obté un conjunt de valors discrets, on cada valor discret representa un rang continu de valors

Atribut continu



Atribut discret





Mètodes de discretització

- Intervals uniformes múltiples
- Intervals definits per l'usuari
- Class-Attribute Interdependence Maximization CAIM [Kurgan and Cios, 2001]
- One Rule Discretizer [Holte, 1993]
- ChiMerge [Kerber, 1992]
- Partició recursiva d'entropia mínima [Fayyad and Irani, 1992]
- Discretització basada en Boxplot [Gibert & R-Roda, 2000]



Valors que falten ("Missing Values")

- **Tècniques d'eliminació**
 - Per eliminar casos amb valors que falten (eliminar files de la Biblioteca de Casos)
 - No es té en compte en el càlcul de similitud els atributs que els hi manquen algun valor (eliminar columnes de la Biblioteca de Casos)
- **Tècniques de substitució o imputació (missing values aleatoris)**
 - Substituir els valors que falten per un valor calculat mitjançant una funció heurística (imputació de valors)
 - ♦ Mitjana / Moda de l'atribut
 - ♦ Mitjana / Moda de l'atribut dins d'una finestra temporal/espacial de longitud i
 - ♦ Predictor CBR/Classificador CBR
 - ♦ Una funció heurística específica
- **Tècniques "mandroses" ("lazy")**
 - Si es desconeix un dels valors comparats, la "dissimilitud" entre els valors és màxima (1)
 - Si es desconeixen els dos valors comparats, la "dissimilitud" entre els valors és mínima (0)



Valoració de la similitud

- **Similitud**

- Calculada entre descripcions de casos
- Es podria calcular com:
 - El resultat d'un procés de concordança de característiques coincidents
 - Indirectament a través del concepte contrari: **dissimilaritat** calculada com a funció de distància
- Pot ser dependent del domini

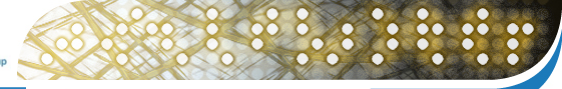
- **Un exemple**

- Descripció del cas = vector d'atributs i valors
- Mesura de similitud $\equiv \text{sim}(C_i, C_j) = 1 - \text{dissim}(C_i, C_j)$

- Mesura de dissimilitud =
$$\text{dist}(C_i, C_j) = \sum_{k=1}^n w_k \times \text{atr_dist}(C_{ik}, C_{jk})$$

- Les dues funcions sim i dist normalitzades sobre $[0, 1]$

- La recuperació hauria d'intentar **maximitzar** la similitud entre el cas actual i el(s) recuperat(s).



Estratègies d'adaptació

Extret de [Sánchez-Marrè, 2022]

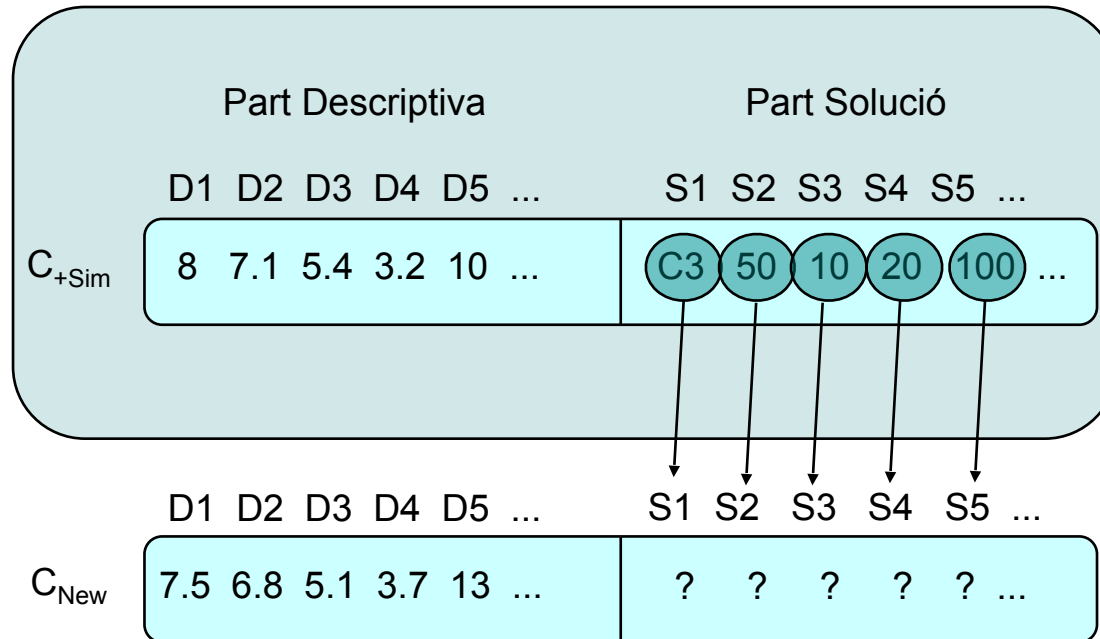
• Adaptació:

- Quan el cas seleccionat de la Biblioteca de casos no coincideix perfectament amb el cas nou, cal que ho sigui la solució antiga **reutilitzada/adaptada** a la solució del nou cas.
- **Estratègies:**
 - ♦ **Adaptació nul·la** (PLEXUS)
 - ♦ **Adaptació ponderada**
 - ♦ **Adaptació estructural**
 - ♦ **Mètodes de substitució**
 - ❖ **Ajustament de paràmetres** (HYPO, PERSUADER, JUDGE)
 - ❖ **Mètodes d'abstracció/re-especialització** (CHEF, JULIANA, SWALE, PLEXUS, CYRUS)
 - ❖ **Mètodes de substitució basats en casos** (CLAVIER, JULIA, CELIA)
 - ♦ **Mètodes de transformació** (CASEY, JULIA, KRITIK)
 - ♦ **Adaptació a un propòsit especial** (PERSUADER, CHEF, JULIA)
 - ♦ **Adaptació derivativa** (ARIES, PRODIGY/ANALOGY, JULIA, MEDIATOR)
- L'adaptació depèn molt del domini



Adaptació nul·la

Biblioteca de casos



- Estratègia de còpia de la solució
 - Per a una variable qualitativa:
 - Per a una variable numèrica:

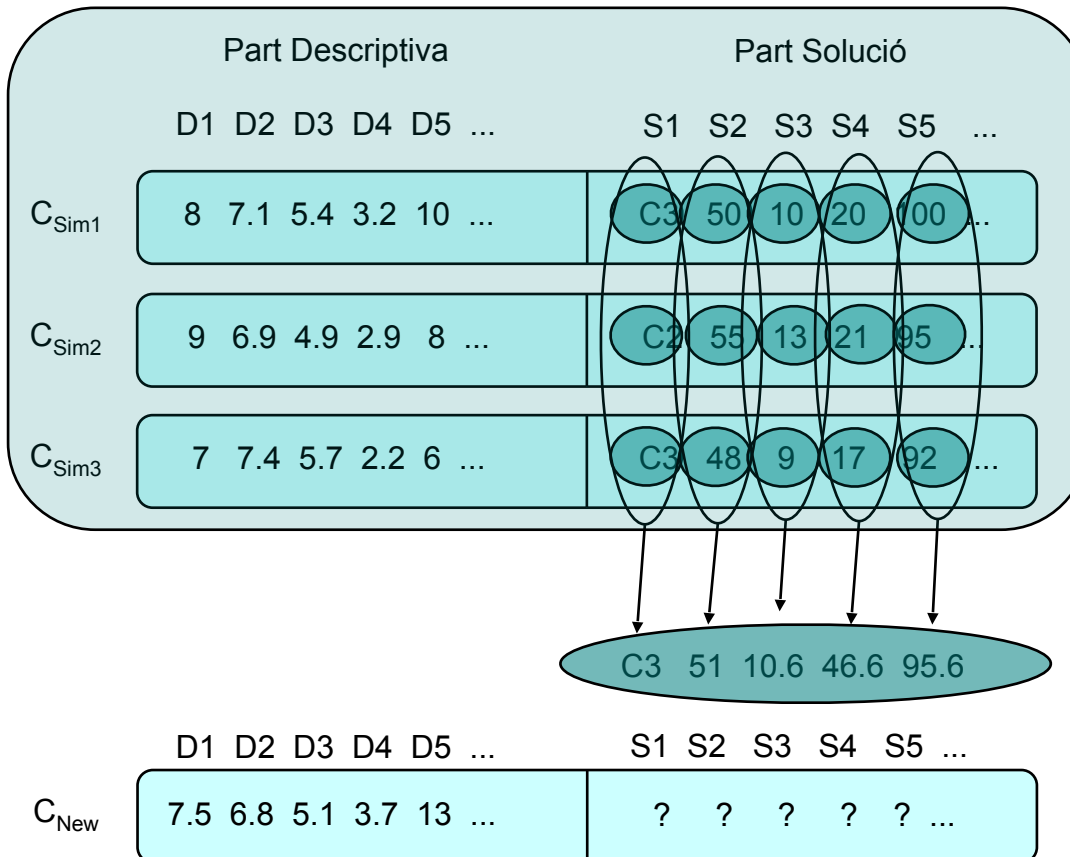
Classificador Nearest Neighbour (NN)!!

Predictor Nearest Neighbour (NN)!!



Adaptació Ponderada

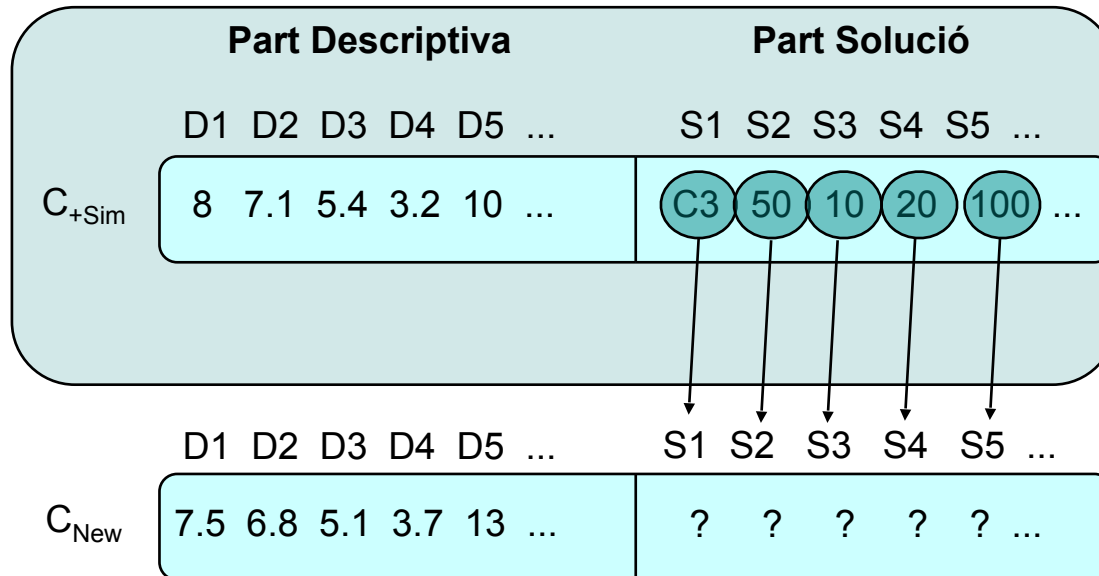
Biblioteca de casos



- Estratègia de solució basada en mitjana/moda
 - Per a una variable qualitativa: **Classificador K-Nearest Neighbour (k-NN)!!**
 - Per a una variables numèrica: **Predictor K-Nearest Neighbour (k-NN)!!**

Estructural / Substitució / Ajustament de Paràmetres

Biblioteca de casos



$$S_i(C_{New}) = f(D_1(C_{+Sim}), ..., D_j(C_{+Sim}), D_1(C_{New}), ..., D_j(C_{New}), S_i(C_{+Sim}))$$

Per ex.,

$$\frac{S_2(C_{+Sim})}{D_1(C_{+Sim})} = \frac{S_2(C_{New})}{D_1(C_{New})} \Rightarrow S_2(C_{New}) = \frac{D_1(C_{New})}{D_1(C_{+Sim})} * S_2(C_{+Sim}) = \frac{7.5}{8} * 50 = 46.875$$

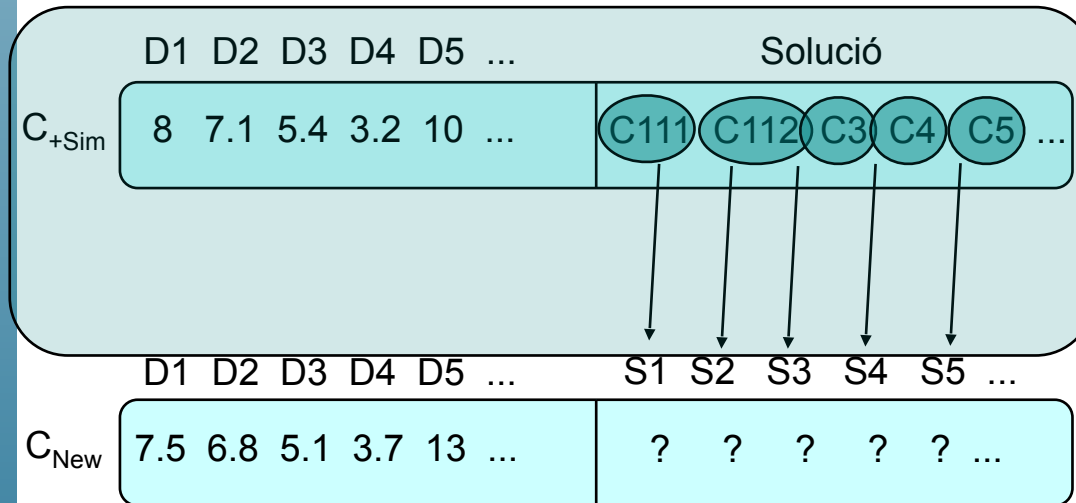


Estructural / Substitució / Abstracció & Re-specialització

Biblioteca de casos

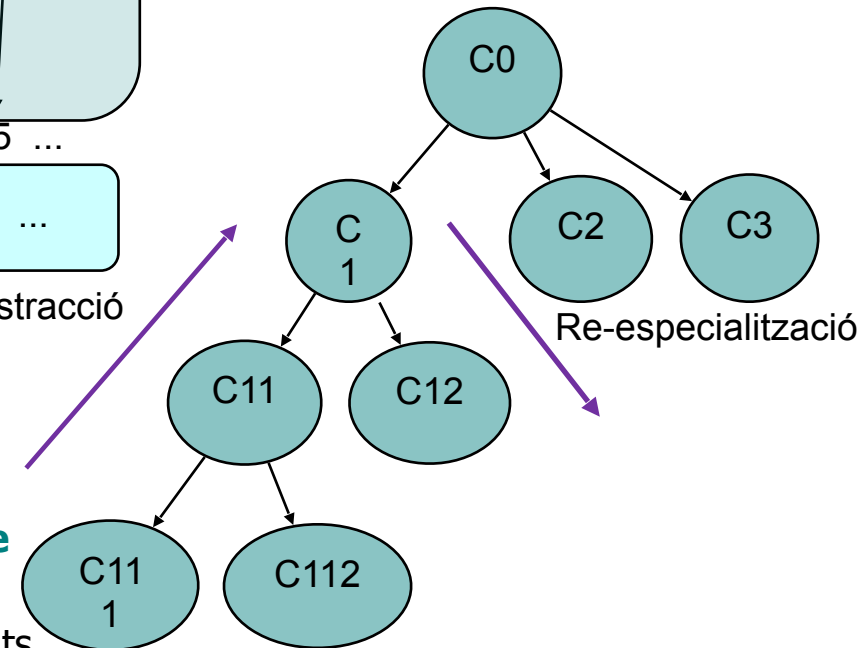
Part Descriptiva

Part Solució



Abstracció

Taxonomia de
Coneixements del
Domini

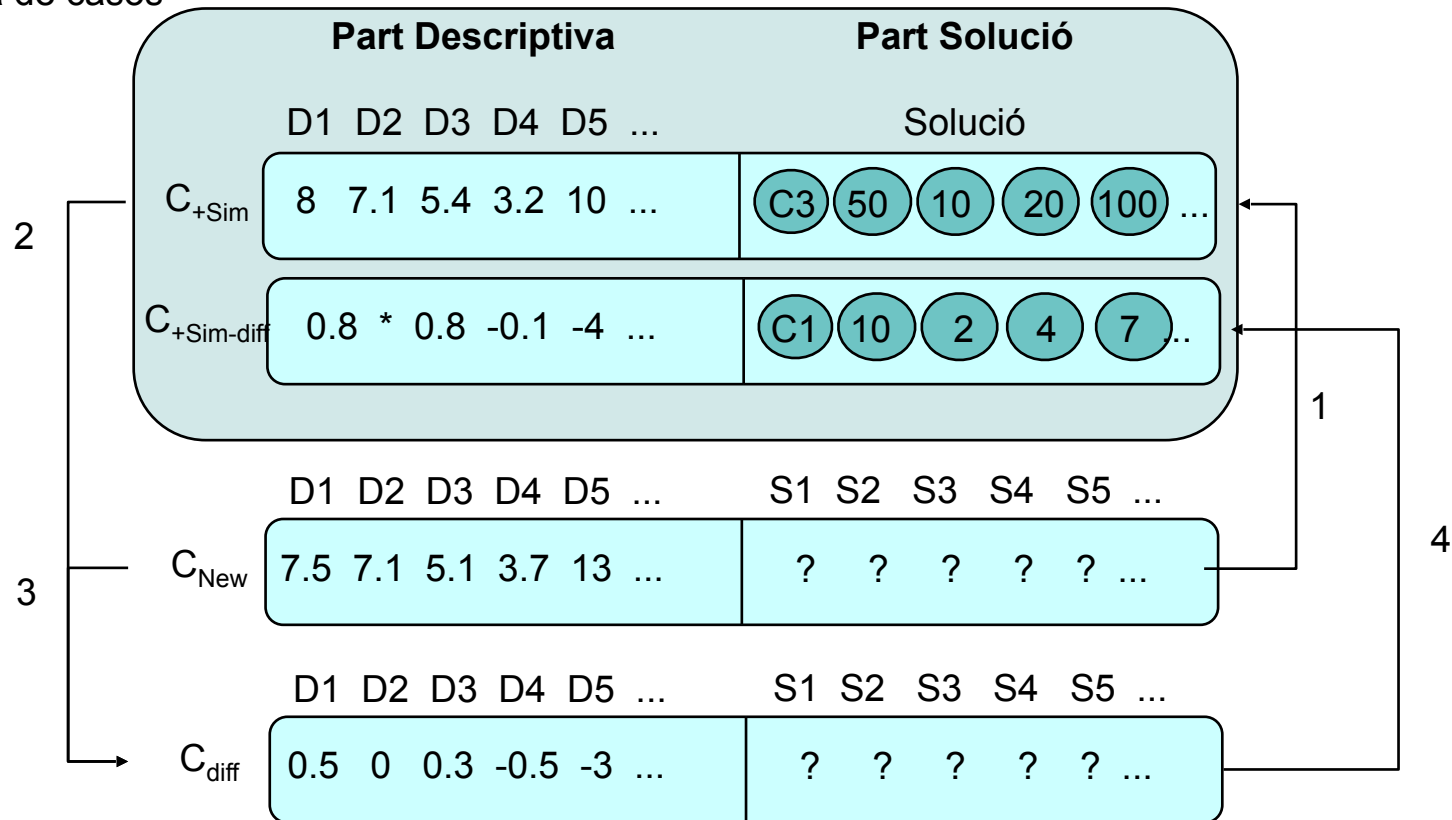


- La solució es compon d'un **conjunt de components**
- S'han de satisfer certes **restriccions**
- Els components que no les satisfaguin **s'han de substituir**
- El component es substitueix per altres components basant-se en una taxonomia/ontologia del coneixement del domini



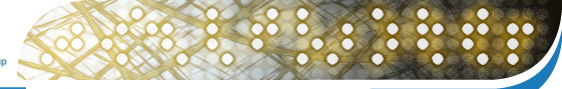
Substitució Estructural i Basada en Casos

Biblioteca de casos



- Recuperar el cas més similar (C_{+Sim}) a C_{new}
- Crear el "cas de diferència" (C_{diff}) entre C_{new} i C_{+Sim}
- Tornar a cercar a la Biblioteca de casos el cas més semblant al C_{diff}
- Utilitzar aquest cas $C_{+Sim-diff}$ més similar a C_{diff}

$$S_i(C_{New}) = f(S_i(C_{+Sim}), S_i(C_{+Sim-diff}), D_1(C_{+Sim}), \dots, D_j(C_{+Sim}), D_j(C_{+Sim-diff}), \dots, D_j(C_{+Sim-diff}), D_1(C_{New}), \dots, D_j(C_{New}))$$



Regles de sentit comú / Estructurals / de Transformació

- La solució s'expressa com un conjunt de components
- Solució = $\{\text{Comp}_1, \text{Comp}_2, \dots, \text{Comp}_n\}$
 - Problema de síntesi (configuració, disseny, planificació, ...)
- **Idea**: aplicar **regles de sentit comú** als components de la solució:
 - Afegir un component
 - Eliminar un component
 - Modificar un component
 - Substituir un component per un altre
- **Exemple**:
 - Una recepta és un conjunt d'ingredients més un pla (conjunt d'accions) per preparar els ingredients de la recepta
 - Una regla d'adaptació de sentit comú podria ser:

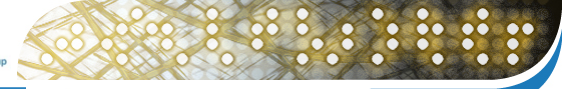
si (si no hi ha mongetes verdes +SimCase_{Filet_de_llom_amb_patates}) **llavors**

Substitueix les patates per mongetes verdes a +SimCase_{Filet_de_llom_amb_patates}

Afegiu un pas d'escalfament dels aliments després d'afegir mongetes verdes

//per eliminar l'excés d'aigua dels aliments

fsi



Adaptacions Estructurals o de propòsit especial

- **Idea:**
- Utilitzar el coneixement basat en experts per fer adaptacions que normalment (CHEF, p.ex.) es coneixen com **Critics repairing rules**
- **Exemple d'arquitectura** (disseny d'edifici, ARCHIE):

if (design of room is incorrect) **then**

if (there is a lack of light) **then**

 set a window in an external wall

elseif (there is not enough space for walking) **then**

 remove a furniture piece

elseif

 ...

endif

endif



Adaptació derivativa/reconstructiva/ per analogia

- **Idea**: no reutilitzar la solució del cas/s més semblant/s, sinó **el mètode utilitzat per derivar /generar la solució**
- Reutilitzar el mètode de solució en el cas nou
- **Exemples**:
 - ARIES (Analogical Reasoning and Inductive Experimentation System) [Carbonell[†], 1983]: **Analogia transformacional**.
 - ♦ El procés de transformació de la solució es va implementar com un conjunt d'operadors de transformació atòmica i un solucionador de problemes de mitjans-fins ("Means-Ends Analysis") que cercava seqüències de transformacions atòmiques i que quan s'aplicaven a la solució recuperada donaven solucions potencials al nou problema. [Carbonell, 1985]
 - PRODIGY System [Carbonell, 1985; Carbonell, 1986; Veloso & Carbonell, 1993]: **Analogia derivacional**
 - ♦ ANALOGY engine a PRODIGY: Un motor d'analogia derivativa [Carbonell & Veloso, 1988] que va poder reproduir solucions senceres a problemes passats similars, cridant al solucionador de problemes de forma recursiva per reduir qualsevol **subobjectiu** nou generat per les diferències conegudes entre els problemes antics i nous.



Avaluació

Extret de [Sánchez-Marrè, 2022]

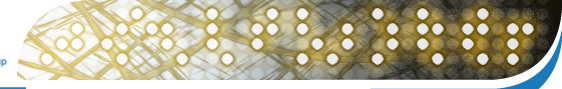
- Avaluació:
 - Valorar el rendiment de la solució proposada
 - Tres formes:
 - ♦ Executar la solució al món real
 - ♦ Preguntar a un expert humà
 - ♦ Executar una simulació



Aprentatge

Extret de [Sánchez-Marrè, 2022]

- Aprentatge:
 - **Aprentatge per observació**
 - ♦ Seeding del casos inicials de la Biblioteca de Casos
 - ♦ Aprentatge d'un nou cas/experiència donat per un expert
 - **Aprentatge per experiència**
 - ♦ Aprentent dels èxits
 - ♦ Aprentent dels fracassos



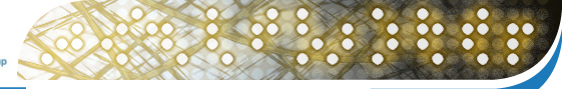
References (1)

- [\[Aha et al., 1999\]](#) D.W. Aha, L.A. Breslow, H. Muñoz-Avila. Conversational Case-Based Reasoning. Applied Intelligence 1-25. 1999.
- [\[Blanzieri & Ricci, 1999\]](#) Blanzieri, E., & Ricci, F. (1999). Probability based metrics for nearest neighbor classification and case-based reasoning. In Proc. of 3rd International Conference on Case-Based Reasoning (ICCBR 99), pp. 14–28. LNAI, Springer.
- [\[Burke et al., 2001\]](#) E.K. Burke, B. MacCarthy, S. Petrovic, and R. Qu. Case-Based Reasoning in Course Timetabling: An Attribute Graph Approach. In Proc. of the 4th International Conference on Case-Based Reasoning (ICCBR 2001), pp. 90-10. Vancouver, BC, Canada, 2001.
- [\[Carbonell, 1983\]](#) Carbonell, J. G. Learning by Analogy: Formulating and Generalizing Plans from Past Experience. In Machine Learning, An Artificial Intelligence Approach (R. S. Michalski, J . G. Carbonell and T. M. Mitchell, eds.), Tioga Press, Palo Alto, CA, 1983.
- [\[Carbonell, 1985; Carbonell, 1986\]](#) Carbonell, Jaime G., Derivational analogy: a theory of reconstructive problem solving and expertise acquisition. Computer Science Department. Paper 1534, 1985. Later published in Machine Learning 2, 1986.



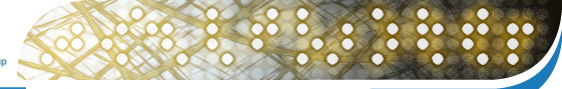
References (2)

- [\[Carbonell & Veloso, 1988\]](#) Carbonell, J.G., & Veloso, M.M. (1988). Integrating derivational analogy into a general problem solving architecture. In Procc. of the 1st Workshop on Case-Based Reasoning (pp. 104-124). Tampa, FL: Morgan Kaufmann.
- [\[Fayyad and Irani, 1992\]](#) U.M. Fayyad and K.B. Irani (1992). On the Handling of Continuous-Valued Attributes in Decision Tree Generation. Machine Learning 8:87-102. 1992
- [\[Gibert & R-Roda, 2001\]](#) Gibert, K. and Roda, I. (2000). Identifying characteristic situations in wastewater treatment plants. In 2nd Workshop on Binding Environmental Sciences and Artificial Intelligence (BESAI 2000) at ECAI 2000, pp. 1–9. Berlin, 2000.
- [\[Golobardes et al., 2000\]](#) Golobardes, E., Llorca, X., & Garrell, J. M. (2000). Genetic Classifier System as a heuristic weighting method for a Case-Based Classifier System. In Proc. of the 3rd Catalan Conference on Artificial Intelligence (CCIA2000).
- [\[Gordon, 1999\]](#) Gordon, A. D. 1999. Classification. 2nd edition Boca Raton, FL: Chapman & Hall/CRC
- [\[Holte, 1993\]](#) R.C. Holte, Very simple classification rules perform well on most commonly used datasets, Machine Learning 11:63-90. 1993.



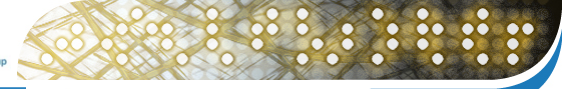
References (3)

- [\[Kerber, 1992\]](#) R. Kerber, ChiMerge: Discretization of Numeric Attributes, In Proc. of the 10th National Conference on Artificial Intelligence (AAAI 92), pp-123-128. MIT Press, 1992.
- [\[Kontkanen et al., 2000\]](#) Kontkanen P., Lathinen J., Myllymäki P. and Tirri H. An unsupervised Bayesian distance measure. In Proc. of 5th European Workshop on Case-Based Reasoning (EWCBR'2000). LNAI-1898, pp. 148-160, 2000.
- [\[Kurgan and Cios, 2001\]](#) L. Kurgan and K. J. Cios. Discretization Algorithm that Uses Class-Attribute Interdependence Maximization. In Proc. of the 2001 International Conference on Artificial Intelligence (IC-AI 2001), pp.980-987, Las Vegas, Nevada.
- [\[Lance & Williams, 1966\]](#) Lance, G. N., & Williams, W. T. (1966). Computer programs for hierarchical polythetic classification ("similarity analyses"). Computer Journal 9:60–64.
- [\[Lenz and Burkhard, 1997\]](#) M. Lenz, and H.D. Burkhard. CBR for Document Retrieval: The FAI/Q Project. In Proc. of the 2nd International Conference on Case-Based Reasoning (ICCBR-97), pp. 84-93. LNAI-1266. Springer. Providence, RI. July 1997.
- [\[Myles and Hand, 1990\]](#) Myles J.P. and Hand D.J. The multi-class metric problem in nearest neighbour discrimination rules. Pattern Recognition 23(11):1291-1297, 1990.



References (4)

- [\[Sánchez-Marrè, 2022\]](#) M. Sánchez-Marrè. Intelligent Decision Support Systems. Springer Nature AG, Switzerland. March 2022.
- [\[Sánchez-Marrè, 1996\]](#) M. Sánchez-Marrè. DAI-DEPUR: an integrated supervisory multi-level architecture for wastewater treatment plants. Ph. D. Thesis. Dept. de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya. 1996.
- [\[Sánchez-Marrè et al., 1998\]](#) Sánchez-Marrè, M., Cortés, U., R-Roda, I., & Poch, M. (1998). L'Eixample distance: A new similarity measure for case retrieval. In Proc. of 1st Catalan Conference on Artificial Intelligence (CCIA'98). ACIA Bulletin 14–15 (pp. 246–253).
- [\[Sanders et al., 1997\]](#) K.E. Sanders, B.P. Kettler, J.A. Hendler. The Case for Graph-Structured Representations. In Proc. of the 2nd International Conference on Case-Based Reasoning (ICCBR-1997), pp. 245-254. LNAI, Springer 1997.
- [\[Short & Fukunaga, 1981\]](#) Short, R. D., & Fukunaga, K. (1981). The optimal distance measure for nearest neighbour classification. IEEE Transactions on Information Theory, 27:622–627.
- [\[Stanfill & Waltz, 1986\]](#) Stanfill, C., & Waltz, D. (1986). Toward memory-based reasoning. Communications of the ACM 29(12):1212–1228.



References (5)

- [\[Veloso and Carbonell, 1993\]](#) M.M. Veloso and J.G. Carbonell. Derivational Analogy in PRODIGY: automating case acquisition, storage and utilization. Machine Learning 10(3):249-278, 1993.
- [\[Weber-Lee et al., 1998\]](#) R. Weber-Lee, A. Martins, and R.M. Barcia. On Legal Texts and Cases. In Proc. of the AAAI-98 Workshop on Textual Case-Based Reasoning. p. 40. 1998.
- [\[Wettschereck et al., 1997\]](#) Wettschereck, D., Aha, D. W., & Mohri, T. (1997). A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artificial Intelligence Review 11:273-314.
- [\[Wilson & Martínez, 1997\]](#) Wilson, D. R., & Martínez, T. R. (1997). Improved heterogeneous distance functions. Journal of Artificial Intelligence Research, 6:1–34.



Intelligent Data Science and Artificial Intelligence (IDEAI-UPC)

Miquel Sànchez-Marrè
miquel@cs.upc.edu



Knowledge Engineering and Machine Learning Group
UNIVERSITAT POLITÈCNICA DE CATALUNYA

<https://kemlg.upc.edu>