



Sciences Economiques et Sociales de la Santé
& Traitement de l'Information Médicale

sesstim.univ-amu.fr

Jean-Baptiste LAMY

*LIMICS (Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en Santé)
Université Sorbonne Paris Nord, Sorbonne Université, INSERM), UMR 1142*

**Intelligence artificielle explicable pour le cancer du sein :
raisonnement à partir de cas visuel**

novembre 2020



Cliquez ici pour voir l'intégralité des ressources associées à ce document

Intelligence artificielle explicable pour le cancer du sein : raisonnement à partir de cas visuel

Jean-Baptiste Lamy



Projet européen H2020
(grant agreement No. 690238)
<http://www.desiree-project.eu>



LIMICS
Université Paris 13, 93017 Bobigny
Sorbonne Universités, Paris
INSERM UMRS 1142

Lamy JB, Sekar B, Guezennec G, Bouaud J, Séroussi B.
Explainable artificial intelligence for breast cancer: a visual
case-based reasoning approach.
Artificial Intelligence in Medicine 2019



Explainable artificial intelligence for breast cancer: A visual case-based
reasoning approach

Jean-Baptiste Lamy^{a,*}, Boomadevi Sekar^b, Gilles Guezennec^a, Jacques Bouaud^{a,c},
Brigitte Séroussi^{a,d}

^a LIMICS, Université Paris 13, Sorbonne Universités, INSERM UMRS 1142, 93017 Bobigny, France

^b School of Computing and Mathematics, Ulster University, United Kingdom

^c AP-HP, DRCI, Paris, France

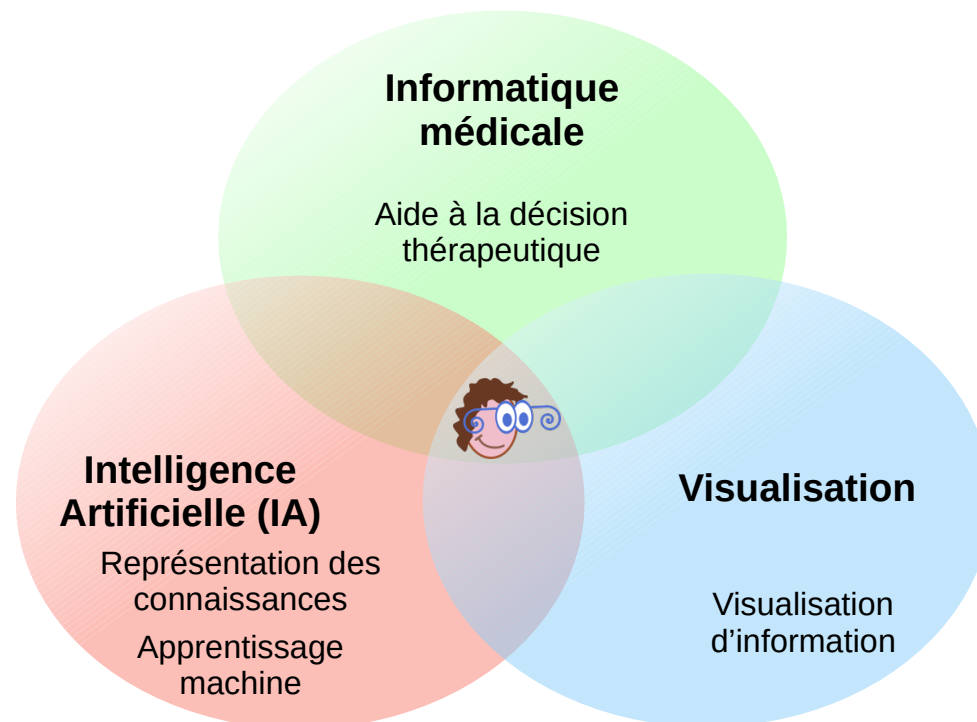
^d AP-HP, Hôpital Tenon, Département de Santé Publique, Paris, France



Présentation

➤ Jean-Baptiste Lamy (PharmD, PhD)

- ◆ Étude de pharmacie puis d'informatique médicale
- ◆ Thématique des recherche :
 - Informatique médicale
 - Aide à la décision thérapeutique
 - Intelligence artificielle (ontologie, apprentissage machine)
 - Visualisation

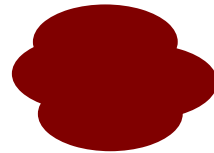


Icônes VCM

➤ VCM : Visualisation des Connaissances Médicales

- ◆ Grammaire combinatoire
- ◆ Ontologie des icônes
- ◆ Alignement avec les terminologies de référence (SNOMED CT)
- ◆ => plus de 3000 icônes

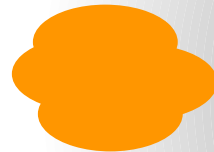
5 couleurs



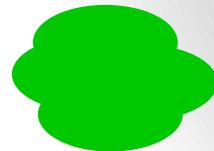
Antécédent



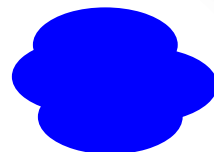
Actuel



Risque



Traitement



Examen

150 pictogrammes centraux



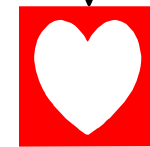
Patient âgé



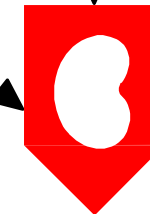
Coeur



Rein



Maladie cardiaque



Insuffisance rénale

Grammaire combinatoire



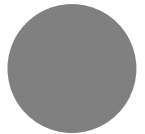
Médicament



Biologie

21 pictogrammes en exposant

35 Formes et modificateurs



État physiologique



Maladie



Hémorragie



Insuffisance



Infection bactérienne

Applications des icônes VCM

◆ Dossiers patients

- Hôpital 
- Ville 
- Suivi de la drépanocytose au Sénégal

Diallo AH, Camara G, Lo M, Diagne I, Lamy JB.
Iconic visualization of sickle cell patients current and past health status.
Studies in health technology and informatics (ICIMTH) 2020;272:71-74

— Analyses

📄 Liste des Résultats

📄 Résultats validés

🔍 LES RESULTATS DES ANALYSES



Prénom
Bébé Oumou

Lieu de naissance
CHRSL

Frise chronologique



Anémie

* Anémie normocytaire hypochrome régénérative avec une hyperleucocytose et une thrombocytose modérée. Au frottis on note la présence de quelques drépanocytes.

Paludisme

* Plasmodium retrouvé dans le sang du patient.

BIOCHIMIE

HEMATOLOGIE

Applications des icônes VCM

◆ Dossiers patients

- Hôpital 
- Ville 
- Suivi de la drépanocytose au Sénégal

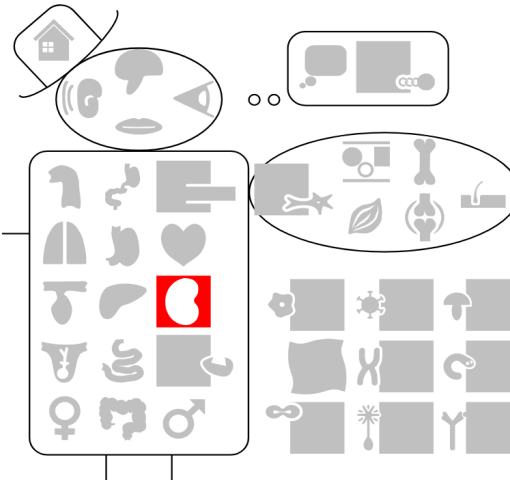
◆ Navigation dans les terminologies médicales

- MedDRA en pharmacovigilance













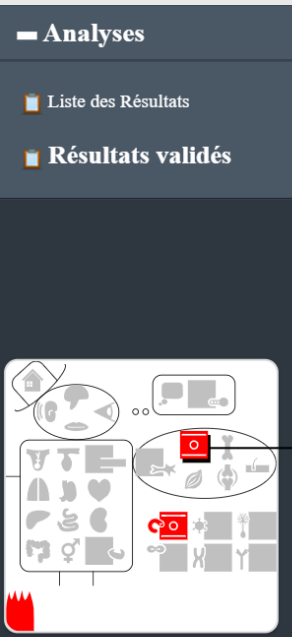
Rechercher dans MedDRA :

abcès



7 termes MedDRA trouvés :

-  Abscès de l'outraque (PT )
Abscès des voies urinaires (PT )
Abscès périnéphrique (PT )
Abscès rénal (PT )
-  Abscès urétrhal (PT )
Abscès à la vessie (PT )
-  Abscès de l'uretère (PT )



Anémie
* Anémie normocytaire hypochrome régénérative avec une thrombocytose modérée. Au frottis on note la présence de drépanocytes.

Paludisme
* Plasmodium retrouvé dans le sang du patient.

Applications des icônes VCM

◆ Dossiers patients

- Hôpital 
- Ville 
- Suivi de la drépanocytose au Sénégal

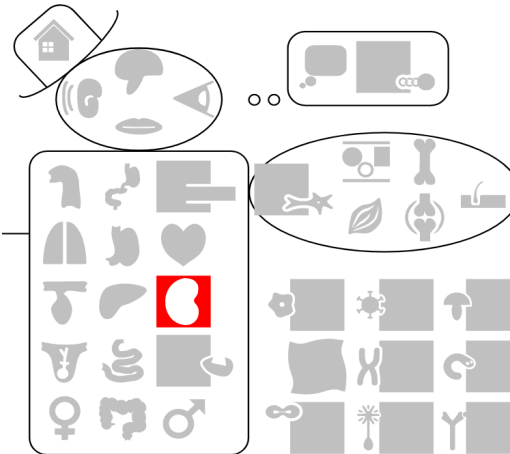
◆ Navigation dans les terminologies médicales

- MedDRA en pharmacovigilance








**MEDINFO
2019
#1 best paper**

Rechercher dans MedDRA :

abcès



7 termes MedDRA trouvés :

-  Abscès de l'ouraque (PT ≡)
-  Abscès des voies urinaires (PT ≡)
-  Abscès périnéphrique (PT ≡)
-  Abscès rénal (PT ≡)
-  Abscès urétral (PT ≡)
-  Abscès à la vessie (PT ≡)
-  Abscès de l'uretère (PT ≡)

◆ Consultation des connaissances

- RCP des médicaments (notices médecin)
- Guide de bonnes pratiques cliniques



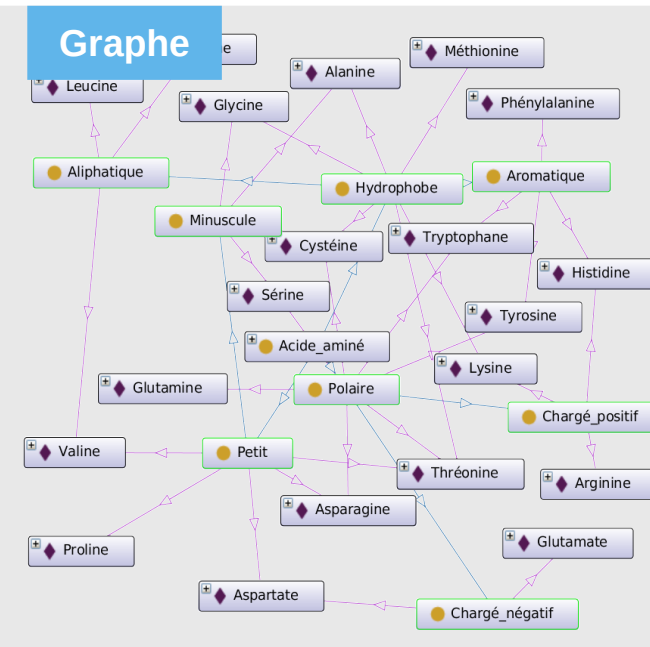
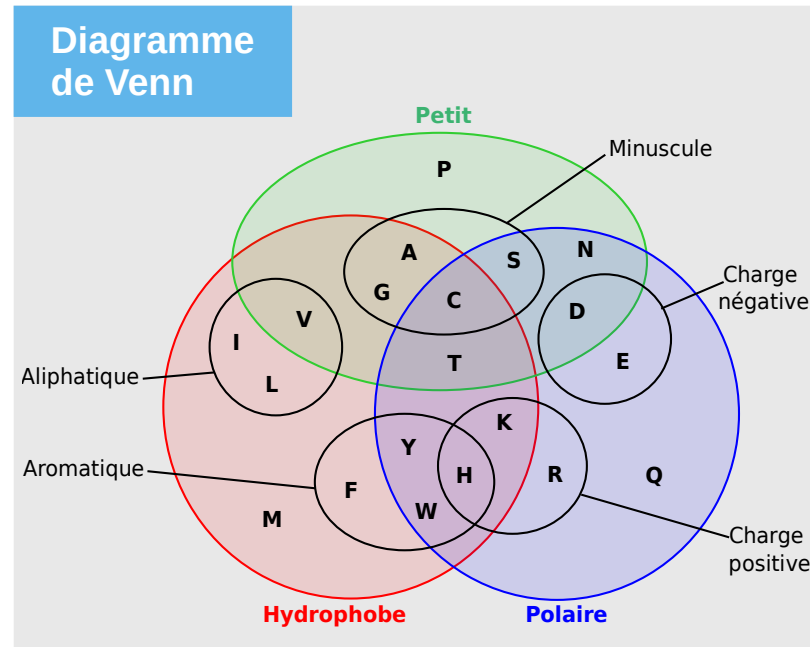
- Moteur de recherche médicaux

Doc'CiSMeF
Outil de recherche en santé

Visualisation d'ensembles

Visualiser des éléments et des ensembles

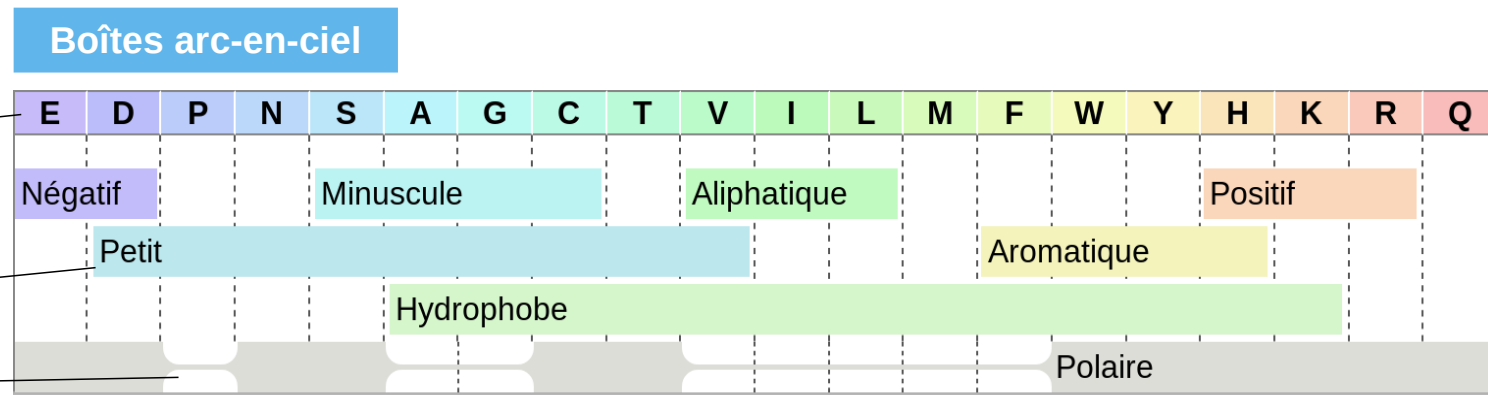
- ◆ Chaque élément appartient à un ou plusieurs ensembles
- ◆ Ex: propriétés des acides aminés
- ◆ Algorithme d'optimisation combinatoire pour ordonner les colonnes



Élément
= colonne

Ensemble
= boîte rectangulaire

Trou



Application de la visualisation d'ensembles aux médicaments

	Prostaglandine E1				Inhibiteurs de la phosphodiesterase de type 5			
	Vitaros	Muse	Edex	Caverject	Viagra	Levitra	Spedra	Cialis
								Antécédents d'insuffisance cardiaque
							Antécédents de troubles du rythme sévères	Troubles du rythme non contrôlés
							Douleur angineuse au cours des rapports sexuels	
Syncope	🚫							
Hypotension orthostatique	🚫							
Balanite		🚫						
Urétrite masculine		🚫						
			Port d'implant pénien	🚫				
Hypersensibilité prostaglandines								
Rapports sexuels sans préservatif avec femme en âge de procréer ou enceinte				🚫				
Prédisposition au priapisme				🚫				
Malformation anatomique du pénis				🚫				
Pathologie contre-indiquant l'activité sexuelle								
Femme								🚫
Âge	18	15	15	15	18	18	18	18

◆ Évaluation avant – après (22 médecins) :

- Change l'avis du médecin sur le médicament **dans 32% des cas**

IA explicable pour le cancer du sein

Introduction

➤ Cancer du sein : l'un des plus fréquents chez la femme

◆ Traitement très complexe

- Chimiothérapie, chirurgie, radiothérapie, hormonothérapie
- Associations possibles, nombreuses protocoles

◆ Décision en Réunion de Concertation Pluridisciplinaire (RCP)

- 3 minutes par patient !

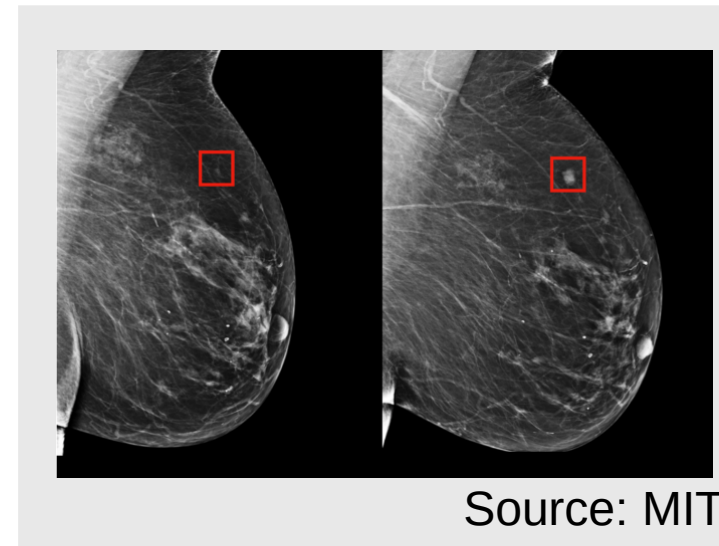
➤ L'intelligence artificielle (IA) peut aider au diagnostic

◆ Analyse d'images

- Explications de l'IA : images annotées

➤ Mais aider à la thérapie est plus complexe !

◆ Comment expliquer le raisonnement de l'IA ?



⇒ IA explicable (*Explainable Artificial Intelligence, XAI*)

IA explicable pour le cancer du sein

Introduction

Projet

- ◆ Projet européen H2020, 11 partenaires
- ◆ **Objectif** : aider à la décision thérapeutique dans le cancer du sein
- ◆ Associe plusieurs méthodes :
 - Aide à la gestion des données patient et des images
 - Implémentation des guides de bonnes pratiques
 - Apprentissage de règles de prescription
 - **Raisonnement à partir de cas**



Introduction

➤ Raisonnement à partir de cas (RAPC, CBR)

◆ Une forme de raisonnement analogique

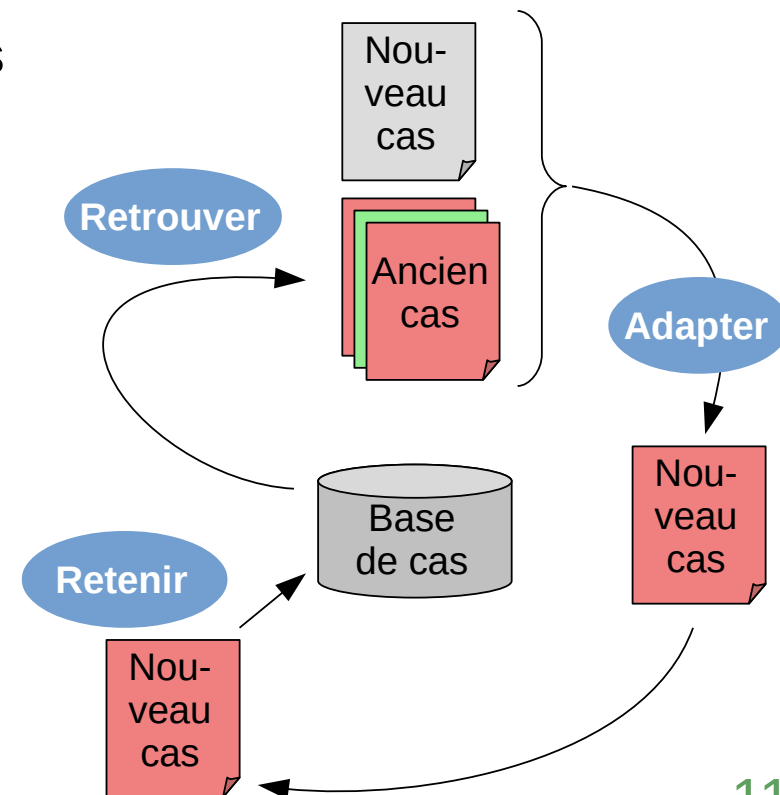
- Exemple type : k plus proche voisin (*k-nearest neighbor*, kNN)

◆ 3 étapes :

- Retrouver les anciens cas similaires dans une base d'anciens cas dont la solution est connue
- Adapter les solutions de ces cas au nouveau cas
- Retenir le nouveau cas dans la base, avec sa solution

◆ Dans le contexte médical :

- Un cas = un patient
- Une solution = un traitement



Introduction

➤ **Raisonnement à partir de cas (RAPC, CBR)**

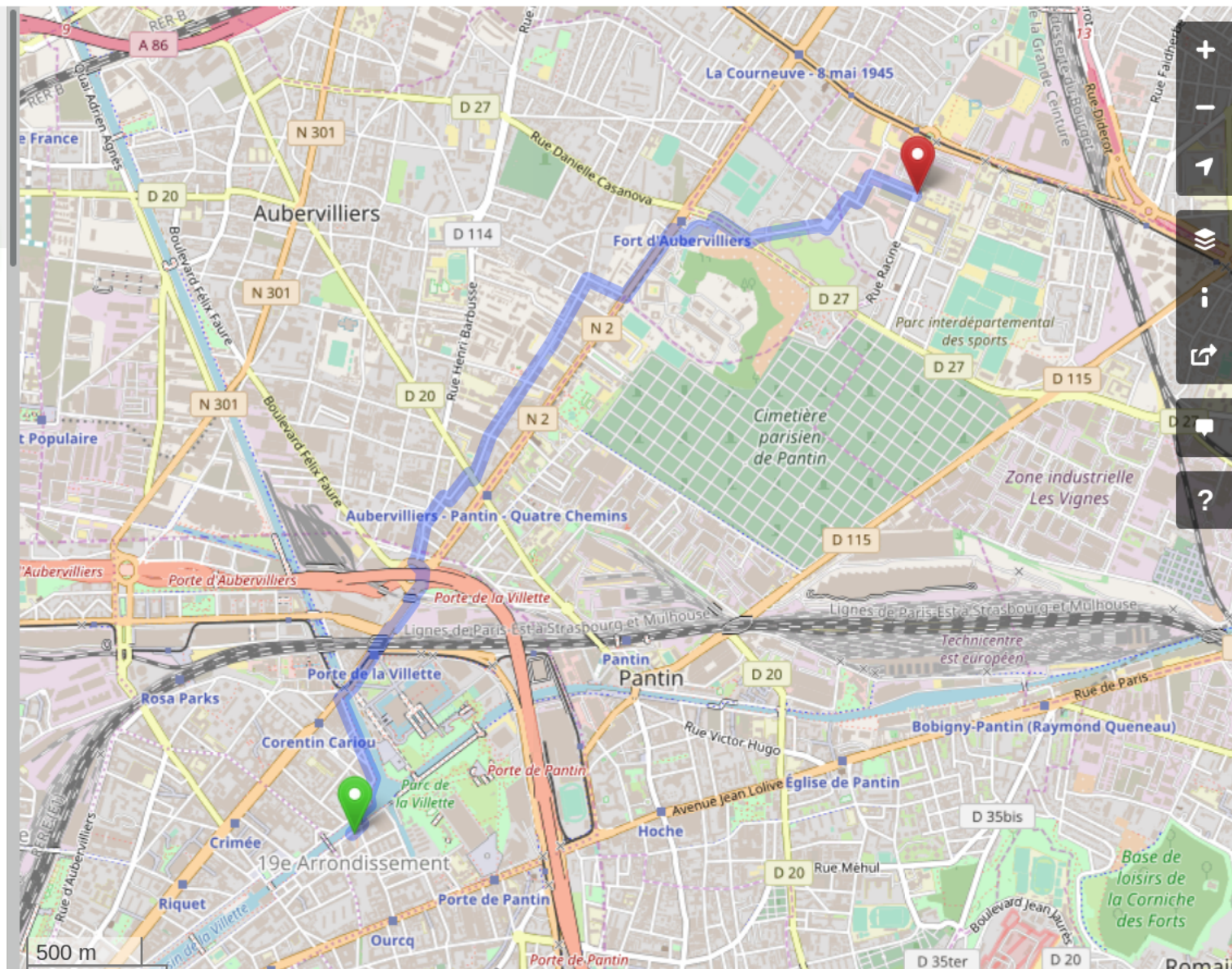
- ◆ IA numérique sans « apprentissage » d'un modèle
 - Le « modèle » consiste en une base de cas passés
- ◆ Très intéressant pour produire des explications :
 - Les anciens cas peuvent servir d'explications
 - Ce mode de raisonnement est familier aux cliniciens

➤ => **Une explication peut consister en la présentation de 5 à 50 cas similaires**

- ◆ Mais 5 à 50 dossiers patient de cancérologie représentent un volume de données très important !
- ◆ Une solution est l'utilisation de la **visualisation d'information** pour présenter les cas et leur similarité avec le nouveau patient

36, Quai de la Marne, Quartier de la Villette
Avenue de la Convention, ZAC Vache à l'Aigle
À pied (GraphHopper)
Aller

- Itinéraire**
- Distance: 4.8km. Temps: 0:58.
Croissant: 23m. Décroissant: 39m.
- ↑ 1. Continuez sur Quai de la Marne 50m
 - ↶ 2. Tournez légèrement à gauche 10m
 - ↷ 3. Tournez légèrement à droite sur Quai de la Marne 10m
 - ↶ 4. Tournez à gauche
 - ↶ 5. Tournez à gauche 70m
 - ↶ 6. Tournez fort à gauche 10m
 - ↶ 7. Tournez à gauche 10m
 - ↷ 8. Tournez à droite sur Quai de l'Oise 500m



Présentation textuelle

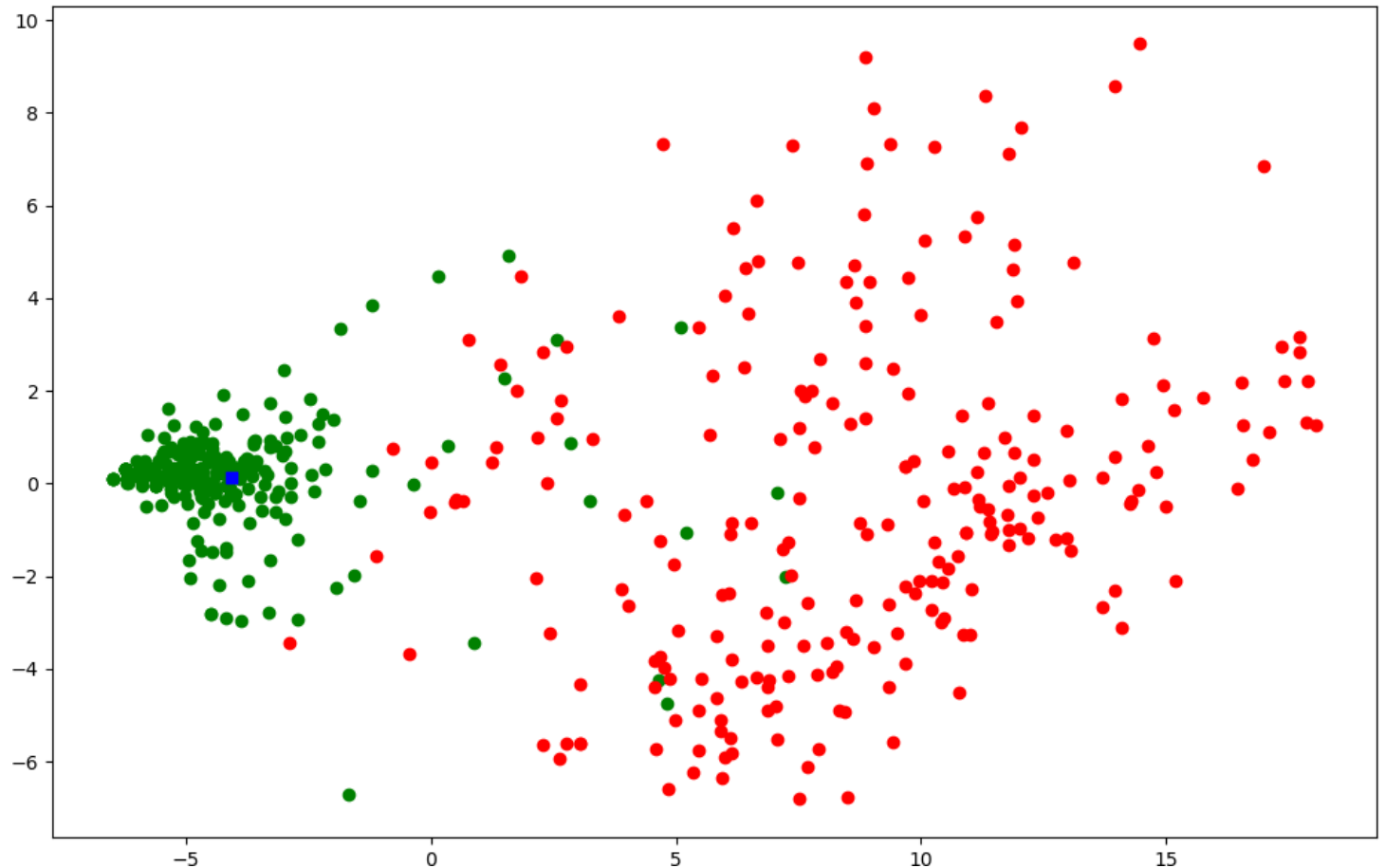
Présentation visuelle

Lecture plus rapide
 Apprentissage et mémorisation
 Modification du trajet possible

Explications textuelles vs visuelles

État de l'art

- Les approches type kNN sont souvent visualisées sous forme de nuage de points



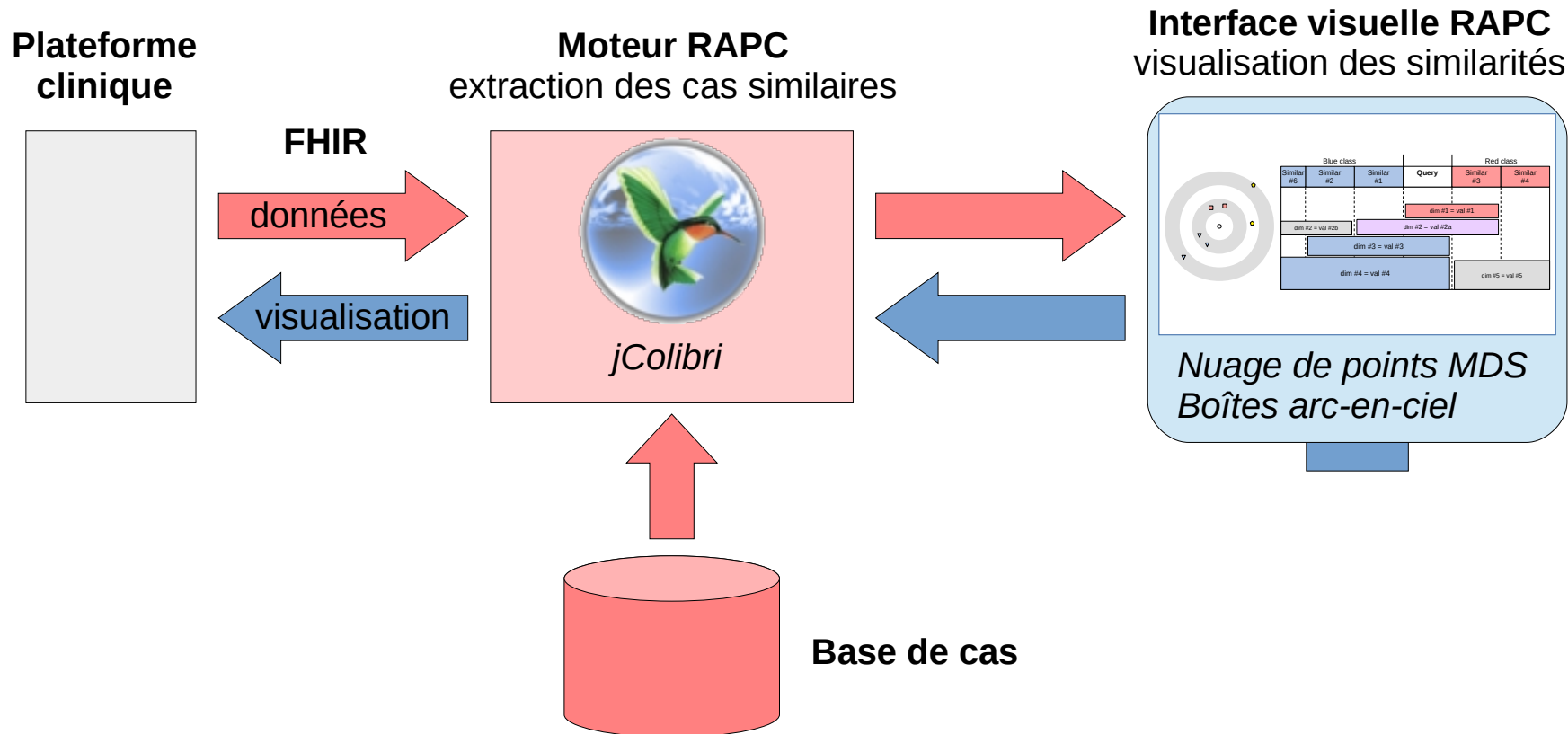
Reste limité : n'explique pas *en quoi* les patients sont similaires !

Objectif

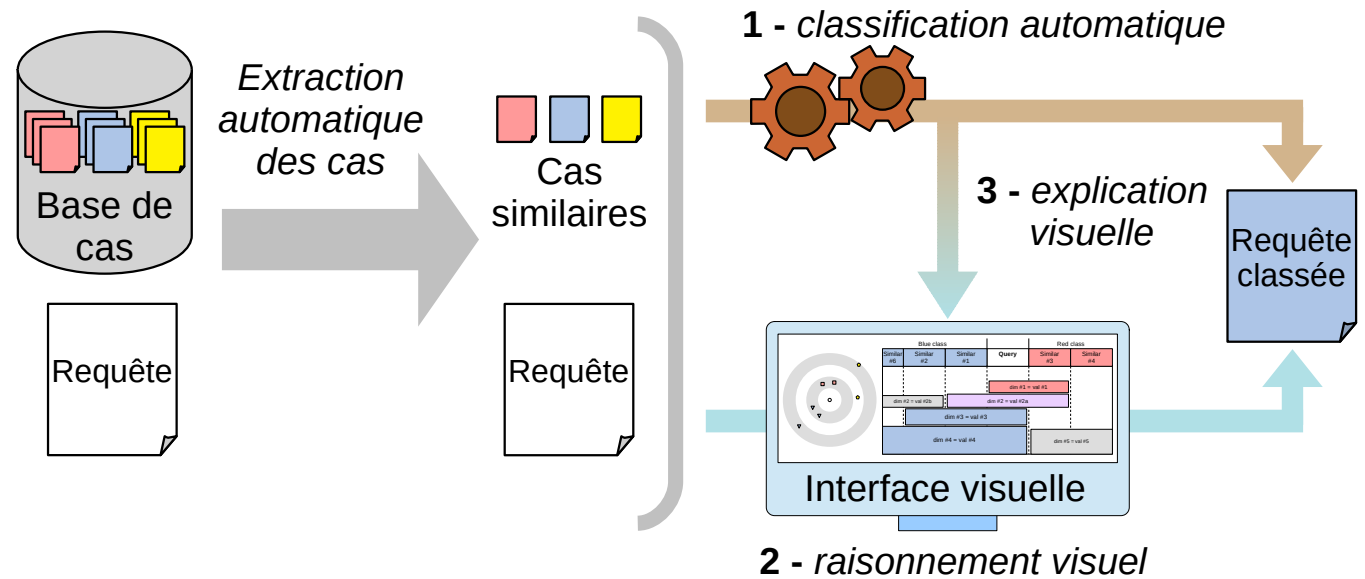
- **Proposer une méthode visuellement explicable pour le raisonnement à partir de cas**
 - ◆ Application au traitement du cancer du sein

Architecture

- ◆ La base de cas est une base de données relationnelle
- ◆ Le standard HL7 FHIR est utilisé pour communiquer avec la plateforme clinique
- ◆ Les cases sont retrouvés avec jColibri



Architecture

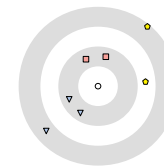


◆ Approche automatique ou visuel

- Traduire visuellement le RAPC

◆ Montre les similarités entre le nouveau cas et les anciens

- Similarités quantitatives
- Similarités qualitatives



Approche quantitative
Montre les similarités
sous forme de distances

Blue class			Query	Red class	
Similar #1	Similar #2	Similar #3		Similar #3	Similar #4
				dim #1 = val #1	
				dim #2 = val #2a	
				dim #3 = val #3	
				dim #4 = val #4	
					dim #5 = val #5

Approche qualitative
Montre les caractéristiques
communes

Raisonnement à partir de cas : phase de recherche des cas

➤ Recherche des n cas les plus proches

➤ Distance euclidienne : $distance(a, b) = \sqrt{\sum_i (attribut\ i(a) - attribut\ i(b))^2}$

◆ Fonctionne bien pour les attributs numériques

- Âge du patient, taille de la tumeur,...

◆ Mais pas pour les attributs booléens / catégoriels

- Localisation de la tumeur, stade de la tumeur, classification TNM,...

➤ Distance sémantique : $dist_{sem}(t_1, t_2) = 1 - \frac{|super(t_1) \cap super(t_2)|}{\sqrt{|super(t_1)|} \times \sqrt{|super(t_2)|}}$

◆ Pour les attributs catégoriels dont les catégories sont les classes d'une ontologie formelle

◆ $super(x)$ est une fonction qui retourne l'ensembles des superclasses d'une catégorie

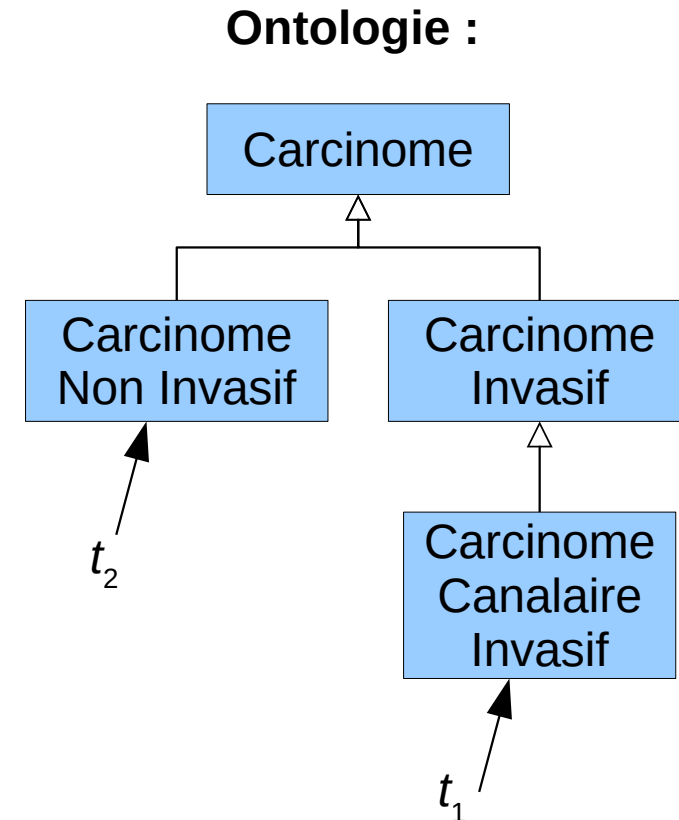
Raisonnement à partir de cas : phase de recherche des cas

$$dist_{sem}(t_1, t_2) = 1 - \frac{|super(t_1) \cap super(t_2)|}{\sqrt{|super(t_1)|} \times \sqrt{|super(t_2)|}}$$

➤ Exemple :

$super(t_1 \in CarcCanalInv) = \{CarcCanalInv, CarcInv, Carc\} \cap \{NonInvCarc, Carc\}$
 $super(t_2 \in CarcNonInv) = \{CarcNonInv, Carc\}$

$$dist_{sem}(t_1 \in CarcCanalInv, t_2 \in NonInvCarc) = 1 - \frac{|\{Carc\}|}{\sqrt{|3|} \times \sqrt{|2|}} = 0.59$$



Utilisation de l'ontologie

➤ Implémentation en Python

➤ Module Owlready 2

- ◆ Programmation orientée-ontologie
- ◆ Permet de charger / enregistrer les ontologies au format OWL
- ◆ Quadstore optimisé (jusqu'à 1 milliard de triplets RDF)
- ◆ Intègre les raisonneurs Pellet et Hermit
- ◆ Accès à UMLS (terminologies médicales)

- ◆ 100 000+ téléchargements

Lamy JB.

Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies.

Artificial Intelligence In Medicine 2017;80:11-28



Python et les ontologies

JB Lamy

ENI éditions, 2019

Matrice des distances

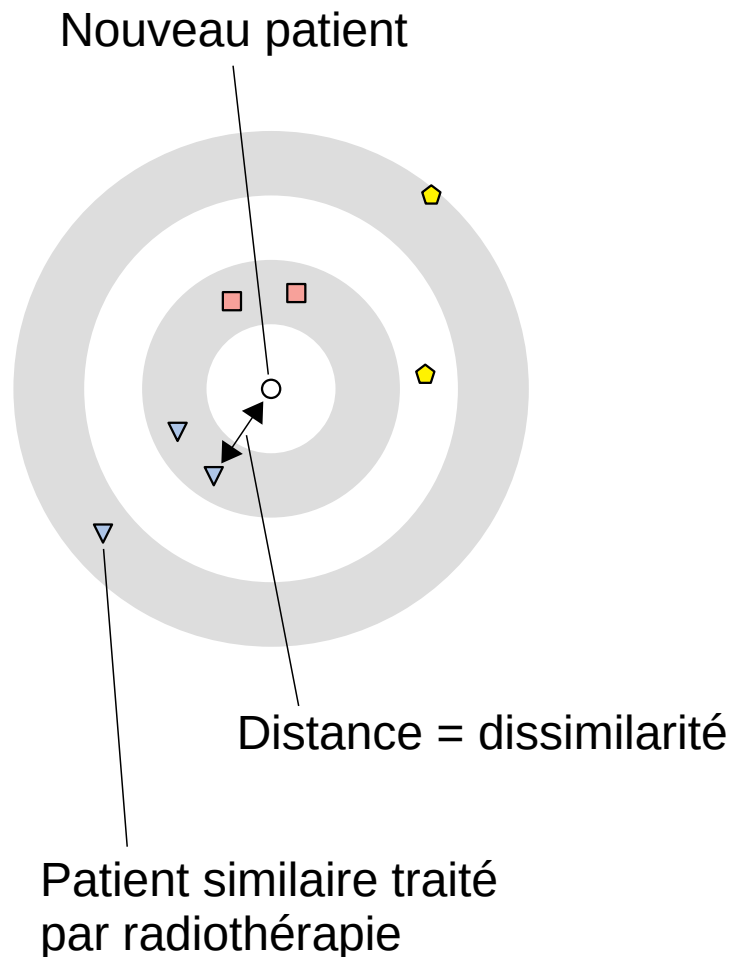
	Query	Similar #1	Similar #2	Similar #3	Similar #4	Similar #5	Similar #6	Similar #7
dim #1	val #1	val #1b	val #1c	val #1	val #1d	val #1e	val #1f	val #1g
dim #2	val #2a	val #2a	val #2b	val #2a	val #2d	val #2e	val #2b	val #2g
dim #3	val #3	val #3	val #3	val #3c	val #3d	val #3e	val #3f	val #3g
dim #4	val #4	val #4	val #4	val #4c	val #4d	val #4e	val #4	val #4g
dim #5	val #5a	val #5b	val #5c	val #5	val #5	val #5e	val #5f	val #5g
dim #6	val #6	val #6a	val #6b	val #6c	val #6d	val #6e	val #6f	val #6g
...								



- ◆ Jcolibri retrouve les cas similaires et calcule la matrice des distances entre cas

	Query	Similar #1	Similar #2	Similar #3	Similar #4	Similar #5	Similar #6	Similar #7
Query	-							
Similar #1	2.0	-						
Similar #2	2.1	1.5	-					
Similar #3	2.0	5.0	4.8	-				
Similar #4	1.9	5.1	4.9	1.1	-			
Similar #5	4.5	5.2	5.2	6.0	6.1	-		
Similar #6	4,2	1.7	1.8	5.5	5.6	5.5	-	
Similar #7	2.0	5.3	5.1	5.4	5.3	3.1	5.2	-

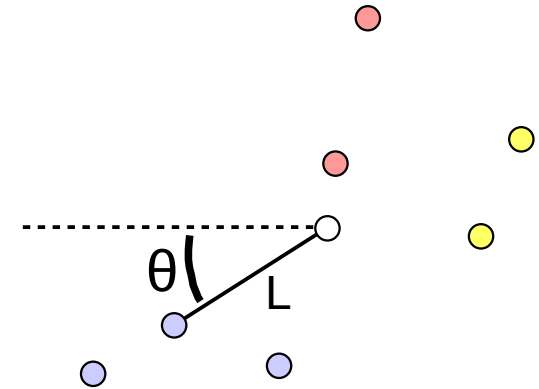
Nuage de points



- **Nuage de points : projection en 2D de la matrice des distances (réduction de dimension)**
 - ◆ 1 point = 1 patient
 - ◆ 1 couleur = 1 classe
 - ◆ La cible au fond facilite l'évaluation des distances
- **Méthodes pour tracer le nuage de points**
 - ◆ MDS (*Multi-Dimensional Scaling*), PCA, tSNE,...
 - ◆ Perte d'une partie de l'information
- **Ici, 2 types de distances :**
 - ◆ A – Entre le nouveau patient et un patient similaire (**les plus importantes !**)
 - ◆ B – Entre deux patients similaires
- **=> Nous proposons une méthode de MDS en coordonnées polaires**
 - ◆ Préserve les distances du type A au détriment des distances de type B

MDS en coordonnées polaires

- ◆ Origine O = le nouveau patient
- ◆ Chaque patient similaire S est défini par ses coordonnées polaires (L, θ)
- ◆ L figure dans la matrice des distances
- ◆ θ est déterminé en résolvant un problème d'optimisation:
 - Trouver les meilleurs valeurs de θ qui minimisent la fonction de stress :



$$S_p(d) = \sum_{2 < i < j} \frac{(d_{ij} - \delta_{ij})^2}{d_{ij}}$$

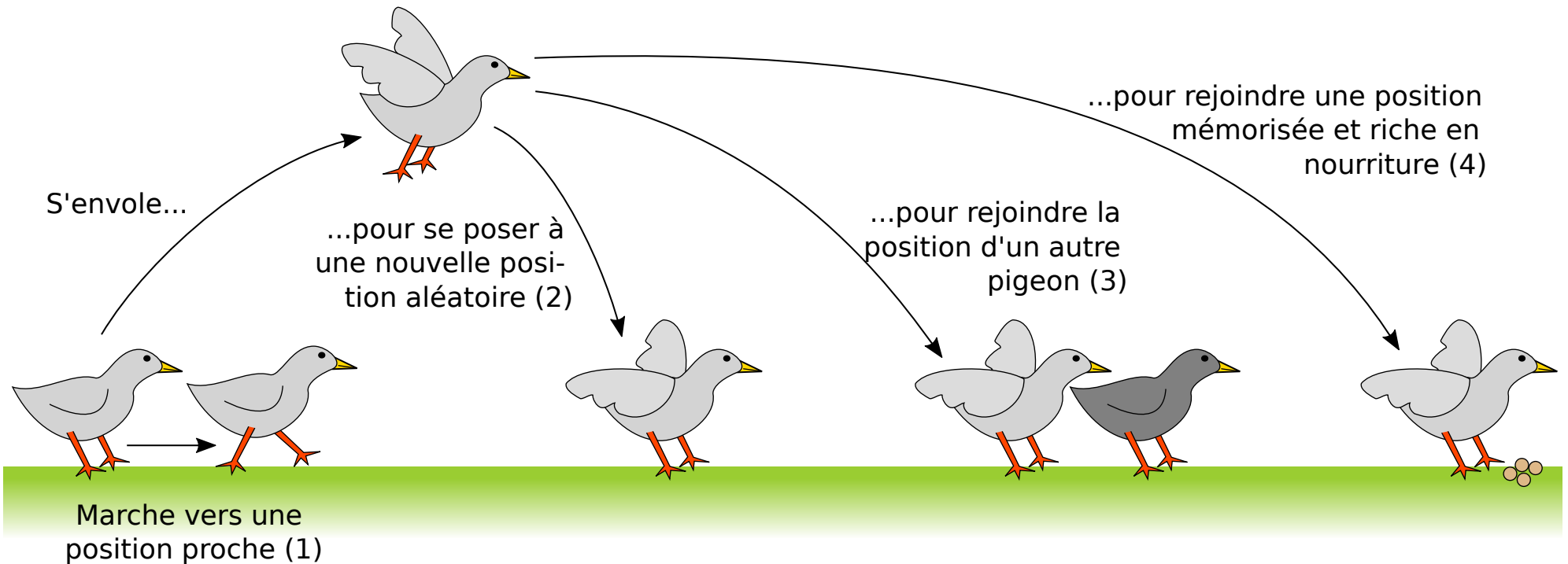
Distance sur le nuage de point en 2 dimensions (dépend des valeurs de θ)

Distance vraie dans n dimensions

Le nouveau patient est le numéro 1

Métaheuristique

Oiseaux artificiels picorant (Artificial Feeding Birds, AFB)
Inspiré du comportement du pigeon



- Simplicité
- Performance
- **Généricité** : optimisation combinatoire, optimisation global non-linéaire,...

Un problème d'optimisation = (*coût()*, *vol()*, *marche()*)

Visualisation des similarités qualitatives

➤ Boîtes arc-en-ciel

- ◆ 1 patient = 1 élément
- ◆ 1 caractéristique partagée = 1 ensemble
 - Ensemble des patients dont l'âge est supérieur à 60,...
- ◆ Seules les deux classes majoritaires sont conservées
- ◆ Les colonnes sont ordonnées par similarité (AFB de nouveau !)
- ◆ Les valeurs numériques sont discrétisées (*Minimum Description Length Principle*, MDLP)
- ◆ Seules les boîtes ayant la plus haute Information Mutuelle (MI) sont conservées

$$MI(ZY) = \sum_{z \in Z} \sum_{y \in Y} p(z, y) \log \left(\frac{p(z, y)}{p(z)p(y)} \right)$$

1)

	Query	Similar #1	Similar #2	Similar #3	Similar #4	Similar #5	Similar #6	Similar #7
dim #1	val #1	val #1b	val #1c	val #1	val #1d	val #1e	val #1f	val #1g
dim #2	val #2a	val #2a	val #2b	val #2a	val #2d	val #2e	val #2b	val #2g
dim #3	val #3	val #3	val #3	val #3c	val #3d	val #3e	val #3f	val #3g
dim #4	val #4	val #4	val #4	val #4c	val #4d	val #4e	val #4	val #4g
dim #5	val #5a	val #5b	val #5c	val #5	val #5	val #5e	val #5f	val #5g
dim #6	val #6	val #6a	val #6b	val #6c	val #6d	val #6e	val #6f	val #6g
...								

2)

	Similar #6	Similar #2	Similar #1	Query	Similar #3	Similar #4
dim #1	val #1f	val #1c	val #1b	val #1	val #1	val #1d
dim #2	val #2b	val #2b	val #2a	val #2a	val #2a	val #2d
dim #3	val #3f	val #3	val #3	val #3	val #3c	val #3d
dim #4	val #4	val #4	val #4	val #4	val #4c	val #4d
dim #5	val #5f	val #5c	val #5b	val #5a	val #5	val #5
dim #6	val #6f	val #6b	val #6a	val #6	val #6c	val #6d
...						

3)

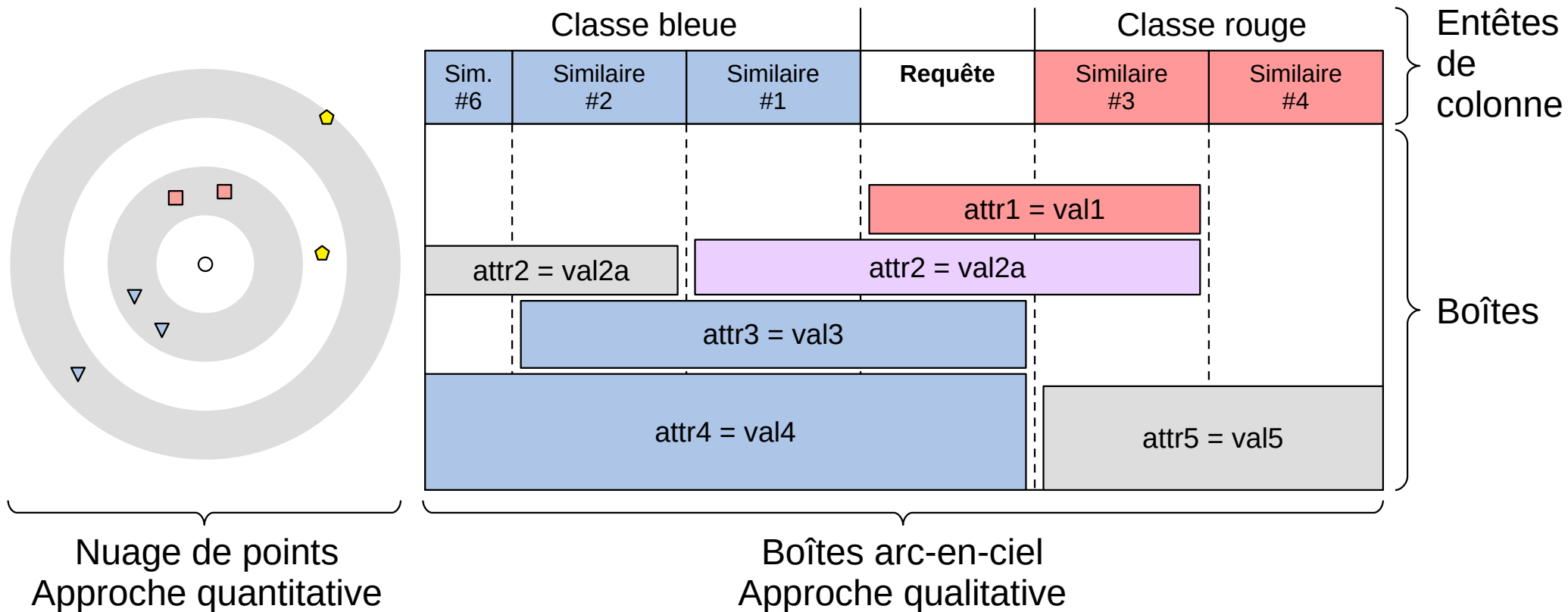
	Similar #6	Similar #2	Similar #1	Query	Similar #3	Similar #4
dim #1	val #1f	val #1c	val #1b	dim #1 = val #1		val #1d
dim #2	dim #2 = val #2b		dim #2 = val #2a			val #2d
dim #3	val #3f	dim #3 = val #3			val #3c	val #3d
dim #4	dim #4 = val #4				val #4c	val #4d
dim #5	val #5f	val #5c	val #5b	val #5a	dim #5 = val #5	
dim #6	val #6f	val #6b	val #6a	val #6	val #6c	val #6d
...						

4)

	Similar #6	Similar #2	Similar #1	Query	Similar #3	Similar #4
				dim #1 = val #1		
	dim #2 = val #2b		dim #2 = val #2a			
	dim #3 = val #3					
	dim #4 = val #4				dim #5 = val #5	

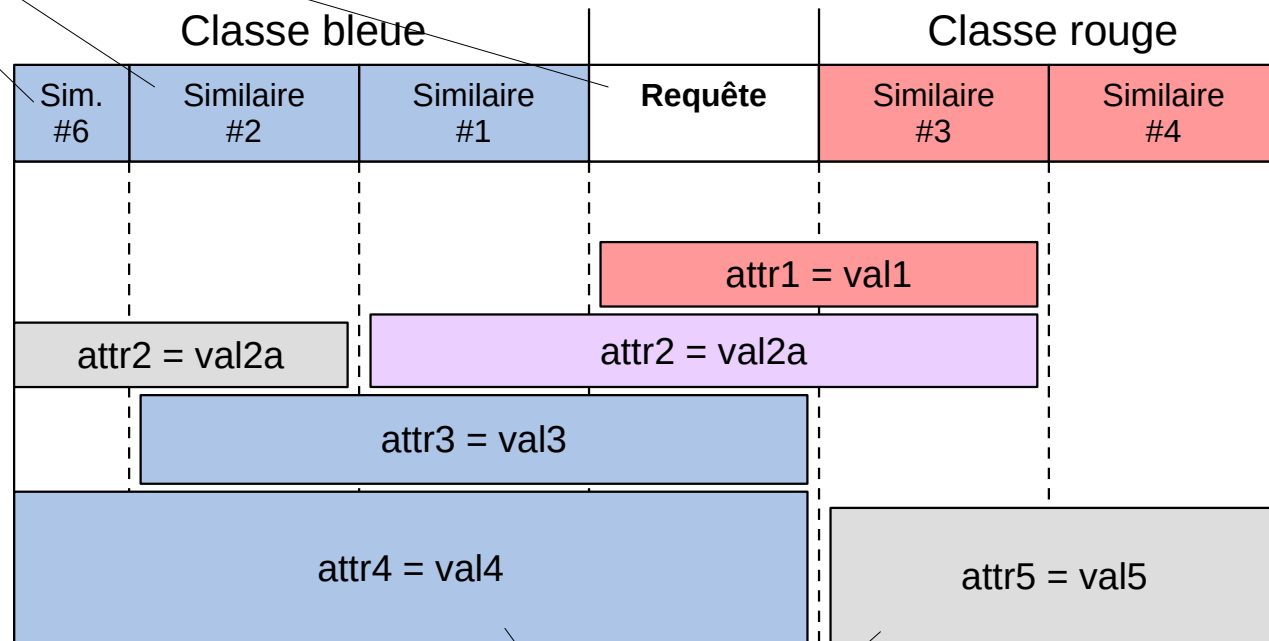
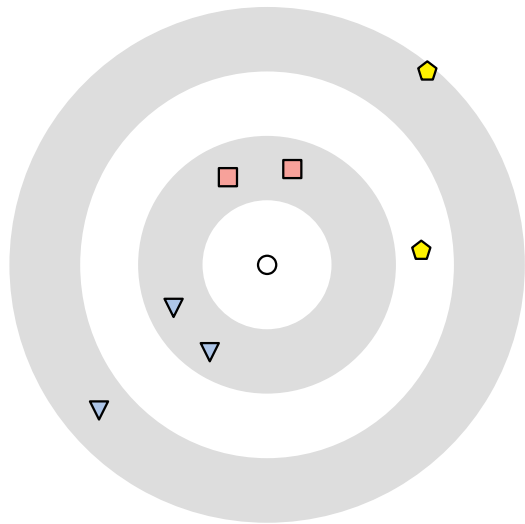
Raisonnement à partir de cas : phase de réutilisation/adaptation

Approche visuelle :



Principes de l'interface

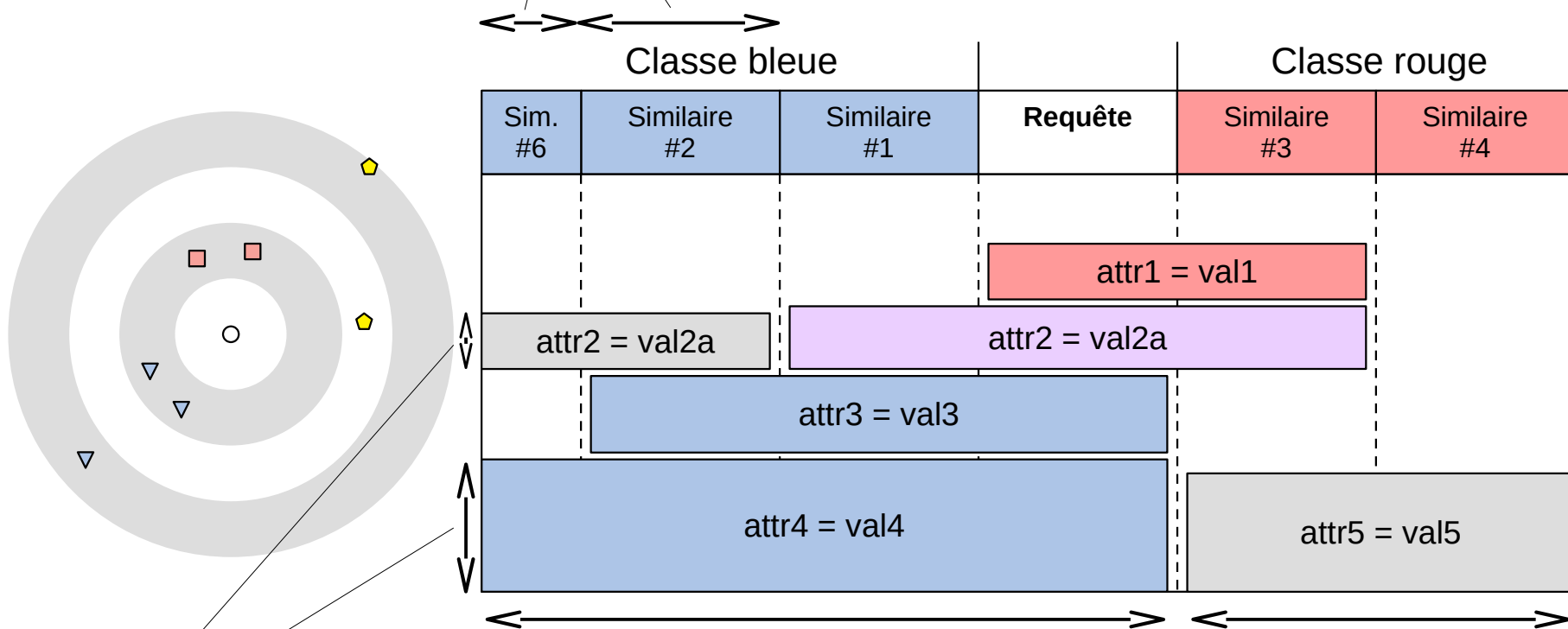
Une colonne dans les boîtes arc-en-ciel = 1 cas



Une boîte = un couple (attribut, valeur)

Principes de l'interface

La largeur des colonnes des cas similaires est proportionnelle à la similarité du cas avec le cas requête

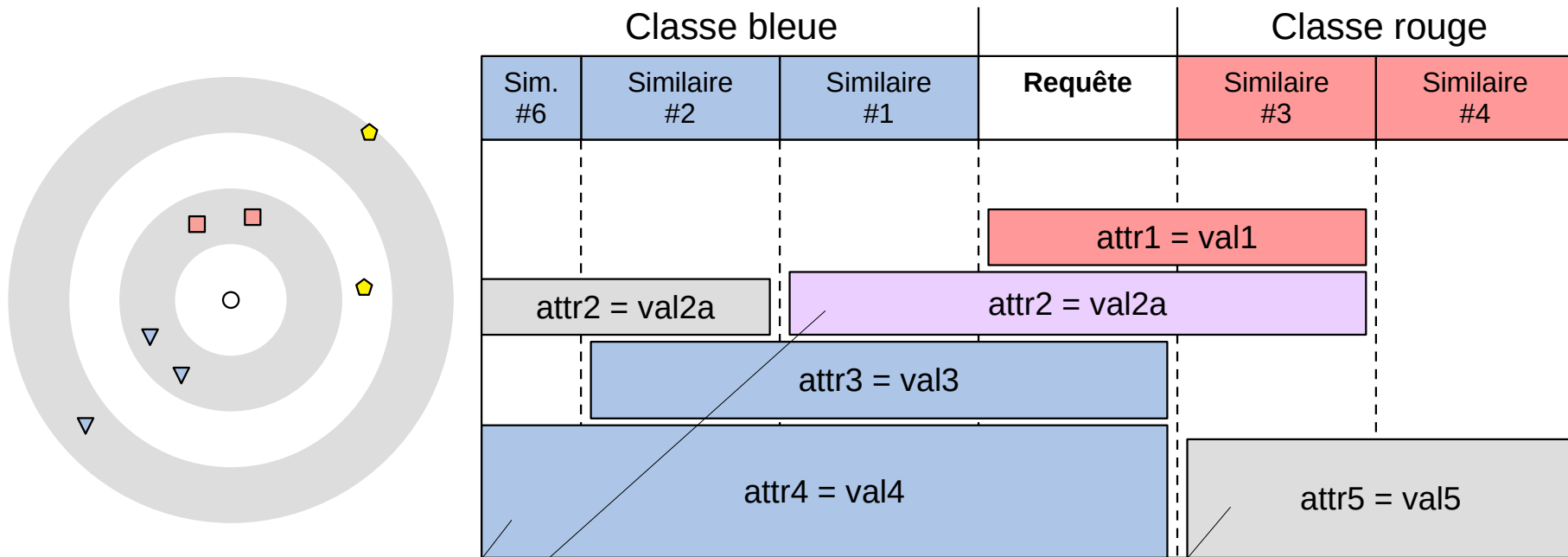


Hauteur des boîtes
= importance
(information mutuelle)

Position horizontale des boîtes
= cas ayant le couple (attribut, valeur)

Principes de l'interface

Les colonnes sont ordonnées de sorte à mettre côte à côte les cas qui partagent des couples (attribut,valeur)

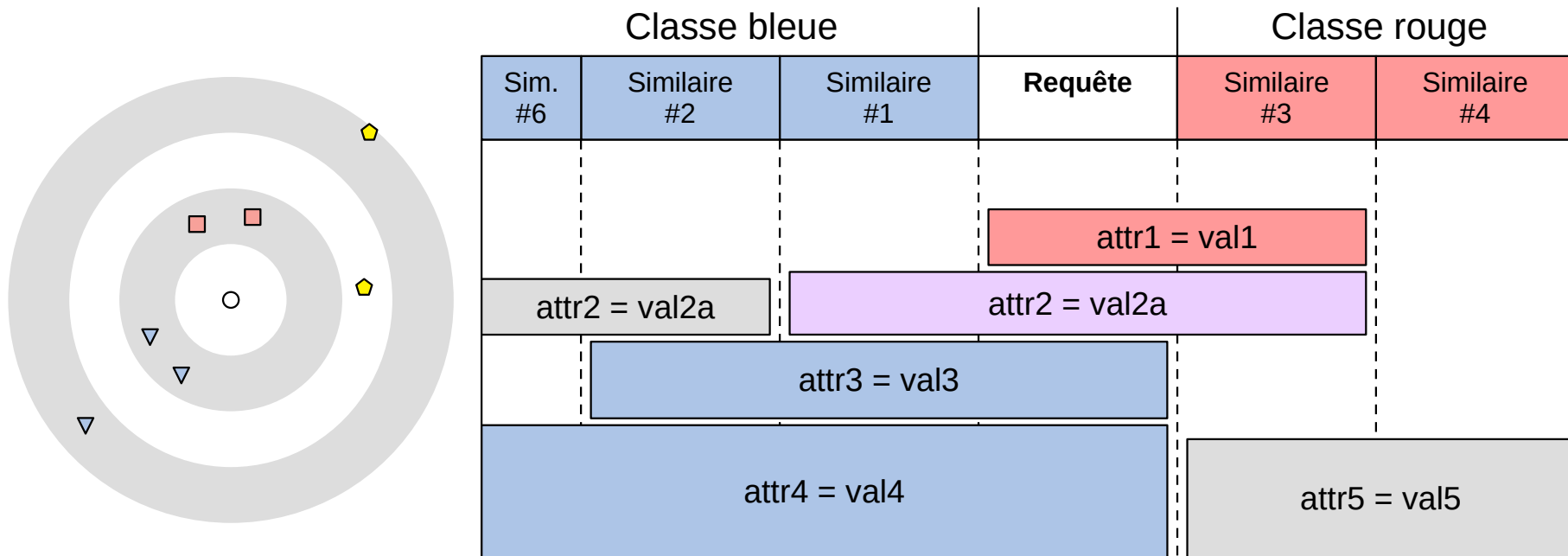


Couleur des boîtes
= moyenne pondérée
des couleurs des colonnes
si la boîte inclut la requête

Couleur des boîtes
= gris si la boîte n'inclut pas la requête

Principes de l'interface

- ◆ Traduit la question « **À quelle classe appartient la requête ?** » en une question visuelle « **Quelle est la couleur dominante ?** »



=> une boîte est un **argument**

Texte = argument qualitatif

Couleur = classe vers laquelle oriente l'argument

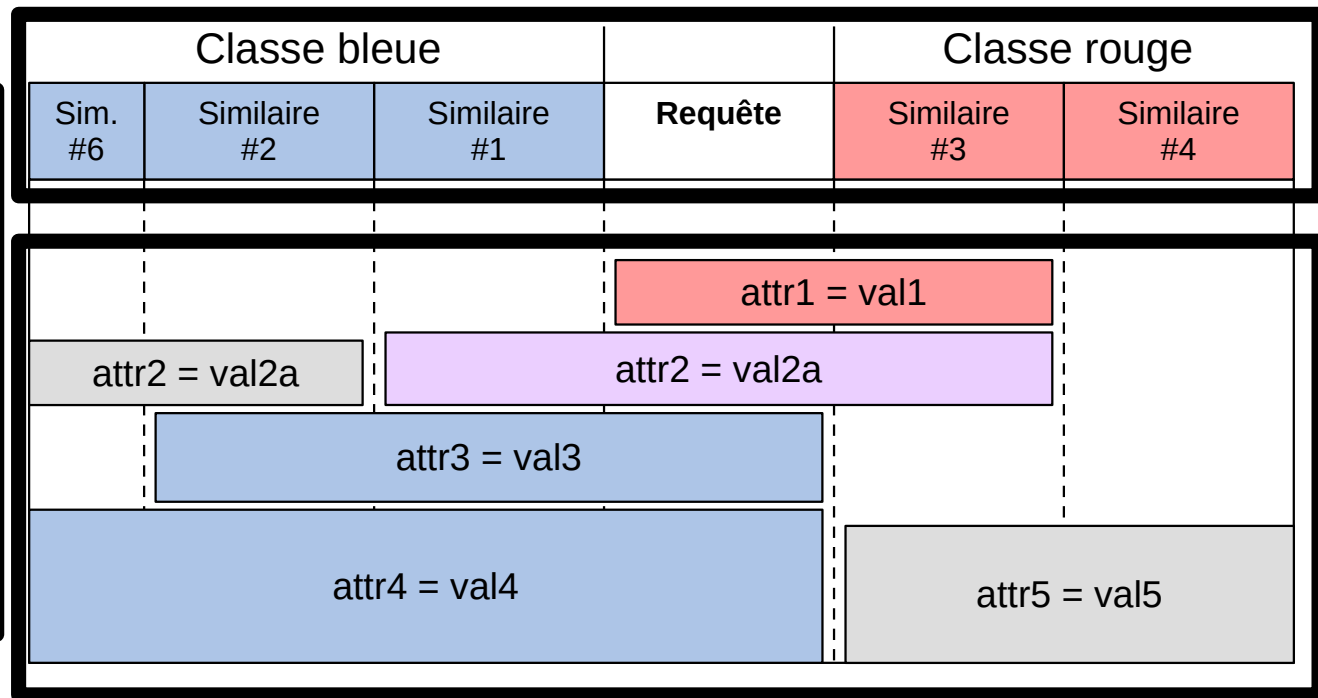
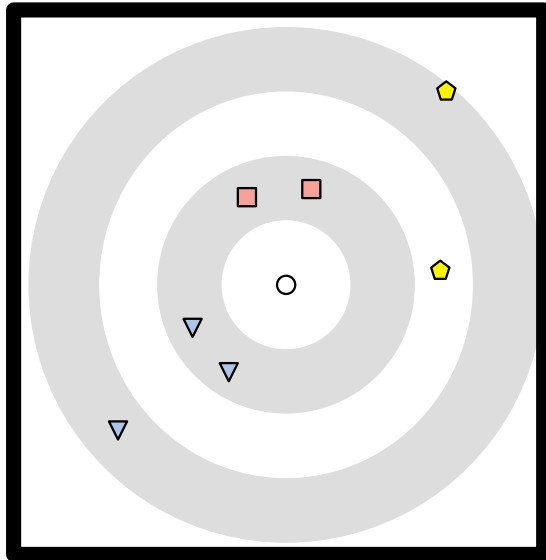
Hauteur = importance de l'argument

Largeur = nombre de cas similaires venant étayer l'argument

Principes de l'interface

kNN / WkNN

WkNN



Nouvel algorithme

La détermination de la couleur majoritaire peut aussi se faire à l'aide d'un algorithme automatique !

« Traduction algorithmique » du raisonnement visuel attendu

function classify(q, X, s, n, m):
 q is the query case
 X is the case database (we assume that $q \in X$)
 s is the dissimilarity measure (a function taking 2 cases and returning their dissimilarity, *e.g.* Euclidean distance)
 $n \geq 2$ is the total number of cases considered (query + similar cases)
 $m \geq 1$ is the maximum number of boxes selected

For each case i in X , compute $s(q, X_i)$
Let X' be the set of selected cases, X' contains the n elements of X with the lowest dissimilarity $s(q, i)$
We assume that $X'_1 = q$ and X'_2 to X'_n are the similar cases

Let d be the distance matrix between cases in X'

For each case i in X' :

For each case j in X' :

$$d_{i,j} = s(i, j)$$

Let w be the weights of the similar cases

$$w_i = \begin{cases} 1 & , \text{if } s_{max} = s_{min} \\ \frac{s_{max} - s(q, X'_i)}{s_{max} - s_{min}} & , \text{otherwise} \end{cases}$$

with $s_{min} = \min(s(q, X'_i))$ and $s_{max} = \max(s(q, X'_i))$

Let y_1 and y_2 be the two best classes in X' (determined by a majority vote over similar cases, weighted by w_i)

Let X'' be the subset of X' displayed in rainbow boxes

$$X'' = X' \cap (\{q\} \cup y_1 \cup y_2)$$

Let B be the set of candidates boxes (currently empty)

For each dimension Z :

If Z has numeric values:

Discretize Z

For each value v that Z takes in cases X'' :

If q has value v for dimension Z :

Add $Zv = \{x \in X'' \mid x_Z = v\}$ **into** B

For each box Zv in B , compute $MI(Z_v Y) = \sum_{z \in \{Z=v, Z \neq v\}} \sum_{y \in \{y_1, y_2\}} p(z, y) \log \left(\frac{p(z, y)}{p(z)p(y)} \right)$

with $p(y) = \frac{|y|}{|X'' \setminus \{q\}|}$, $p(Z = v) = \frac{|Z_v|}{|X'' \setminus \{q\}|}$, $p(Z \neq v) = \frac{|X'' \setminus Z_v|}{|X'' \setminus \{q\}|}$, $p(Z = v, y) = \frac{|Z_v \cap y|}{|X'' \setminus \{q\}|}$, $p(Z \neq v, y) = \frac{|(X'' \setminus Z_v) \cap y|}{|X'' \setminus \{q\}|}$

Let B' be the set of selected boxes, B' contains the m elements of B with the highest $MI(Z_v Y)$

Compute $S_{y_1} = \sum_{Z_v \in B'} \left(MI(Z_v Y) \times \sum \{w_{2 \leq i \leq n} \mid x_i \in Z_v \cap y_1\} \right)$

Compute $S_{y_2} = \sum_{Z_v \in B'} \left(MI(Z_v Y) \times \sum \{w_{2 \leq i \leq n} \mid x_i \in Z_v \cap y_2\} \right)$

If $S_{y_1} > S_{y_2}$:

return y_1

Else:

return y_2

Test sur 3 jeux de données publics

➤ Breast Cancer Wisconsin (BCW)

- ◆ 683 patients (+16 avec des valeurs manquantes, retirés)
- ◆ 9 dimensions entières
- ◆ 2 classes (bénin vs malin)

➤ Mammographic Mass (MM)

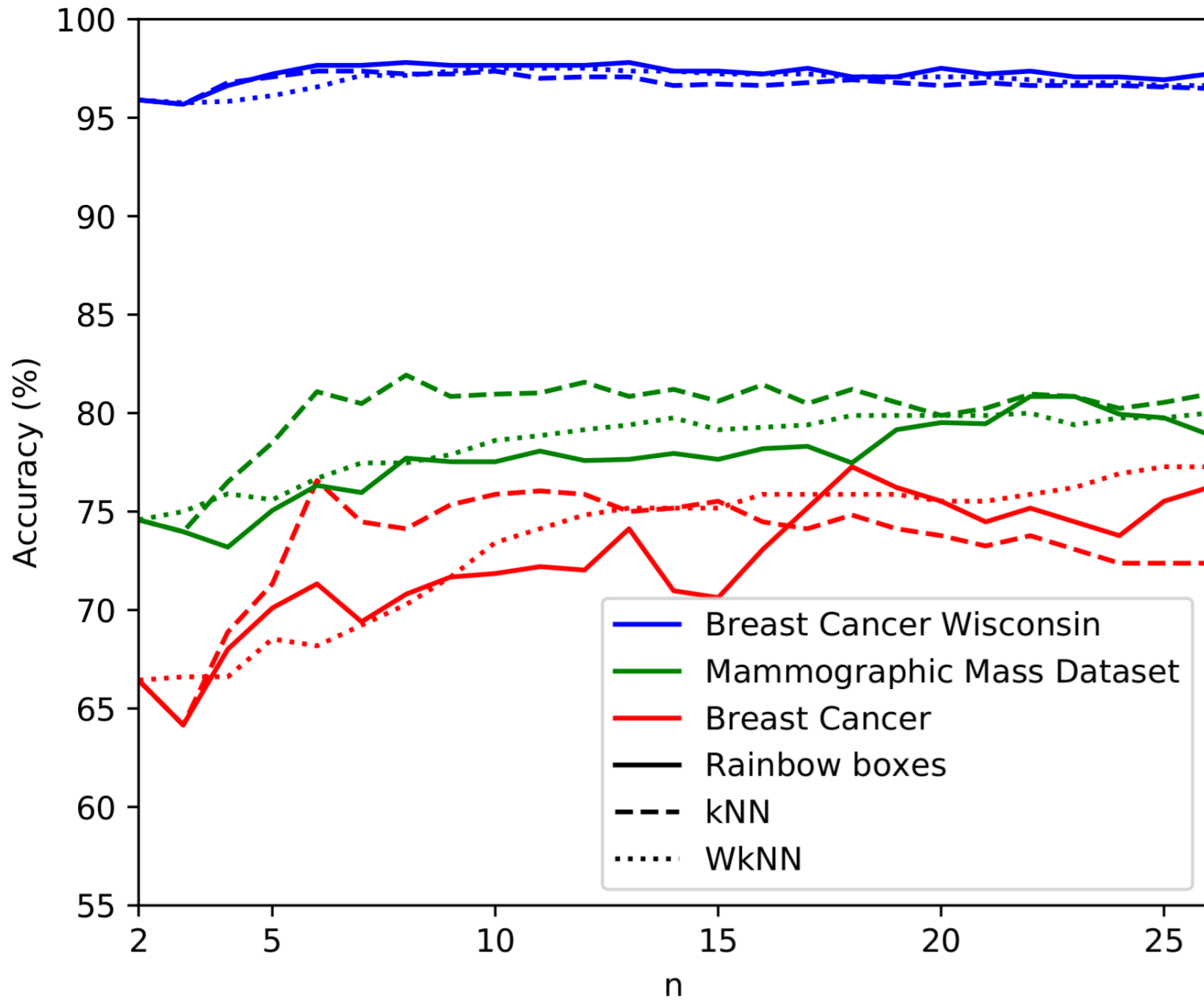
- ◆ 830 patients
- ◆ 2 dimensions entières + 3 dimensions catégorielles
- ◆ 2 classes (bénin vs malin)

➤ Breast Cancer (BC)

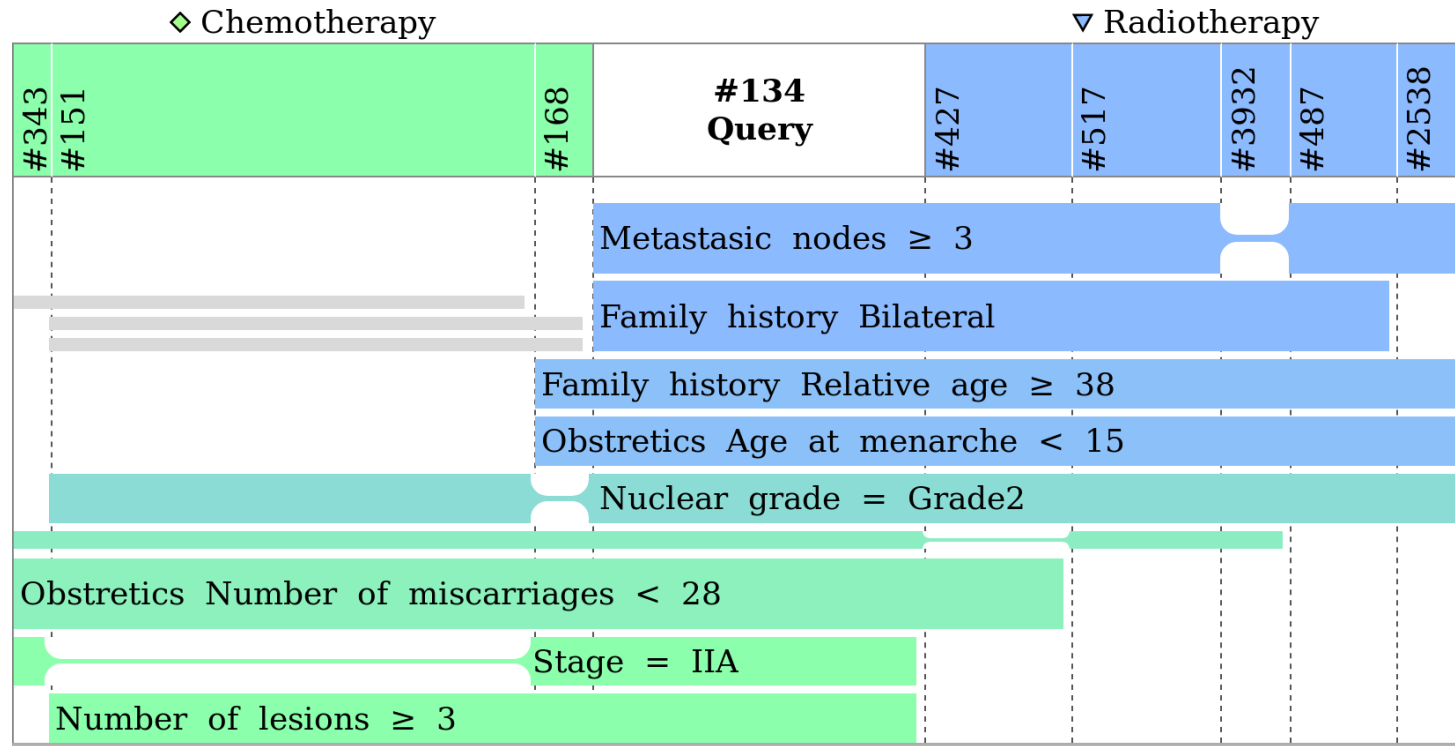
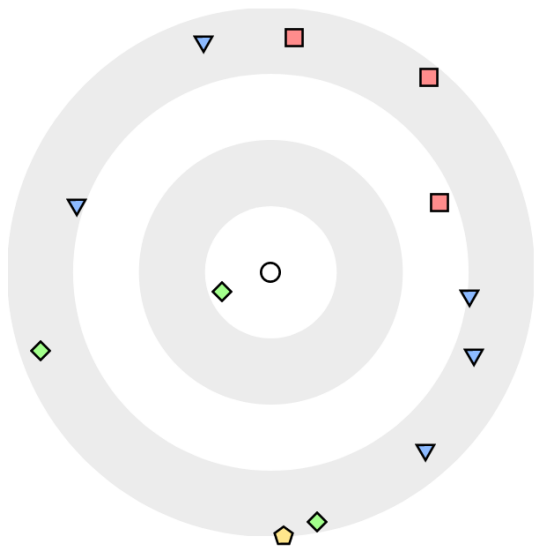
- ◆ 286 patients
- ◆ 4 dimensions numériques + 4 dimensions catégorielles
- ◆ 2 classes (cancer récidivant vs non-récidivant)

Résultats

- ◆ Les éléments visualisés permettent de prendre la bonne décision !
- ◆ Performances équivalentes au kNN
- ◆ Explicabilité meilleure

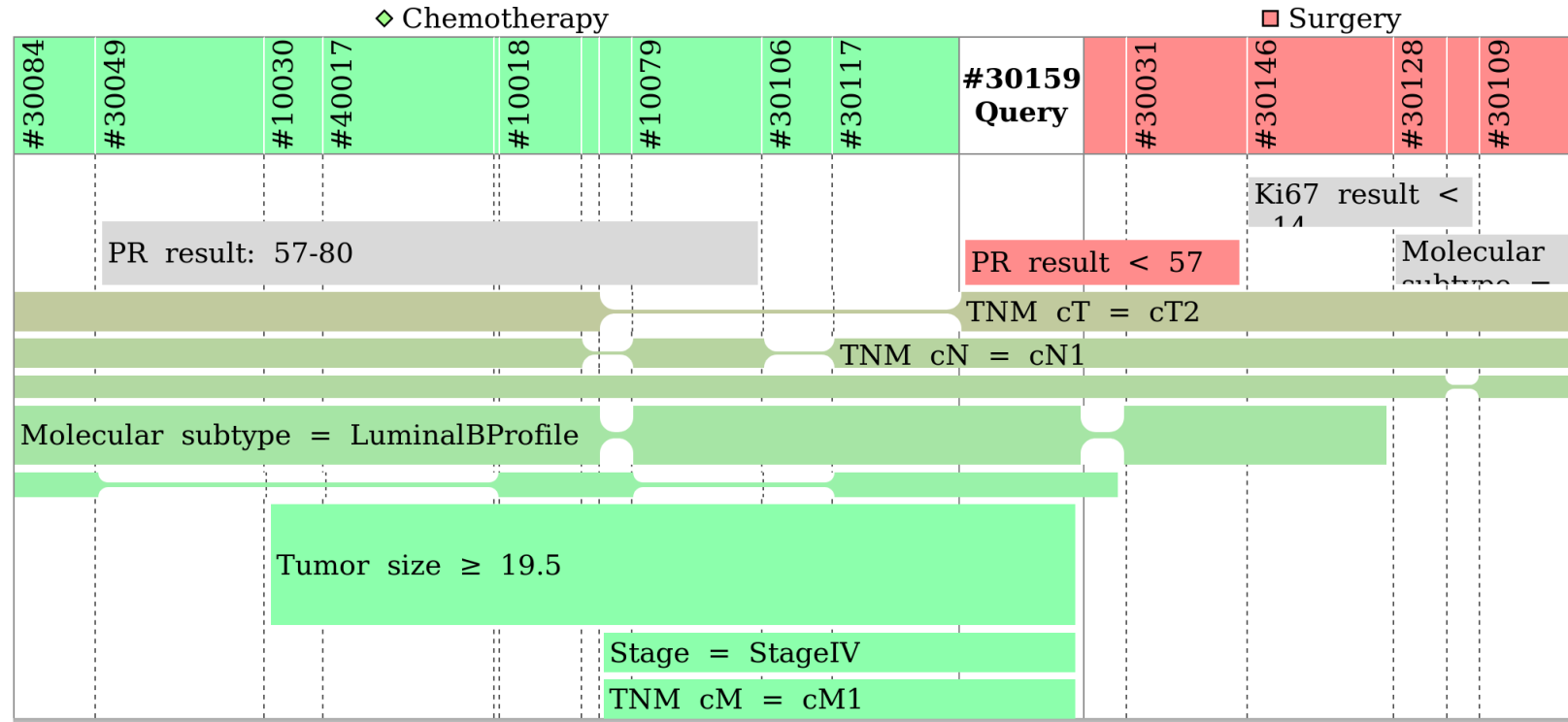
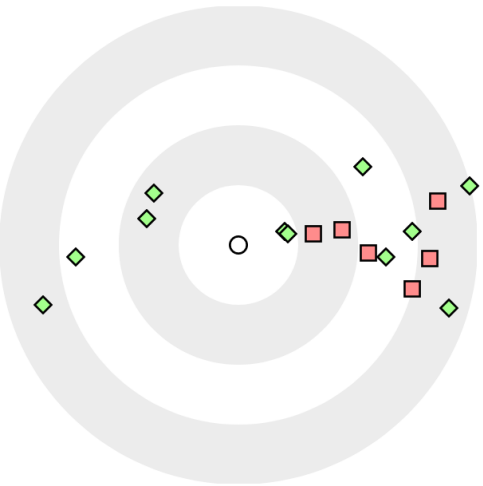


Test sur données réelles anonymisées



- ◆ Les boîtes donnent des arguments en faveur de l'une ou l'autre classes
- ◆ Le traitement recommandé est celui dont la couleur est majoritaire
- ◆ Le clinicien peut choisir une option différente, s'il juge les arguments non pertinents

Test sur données réelles anonymisées



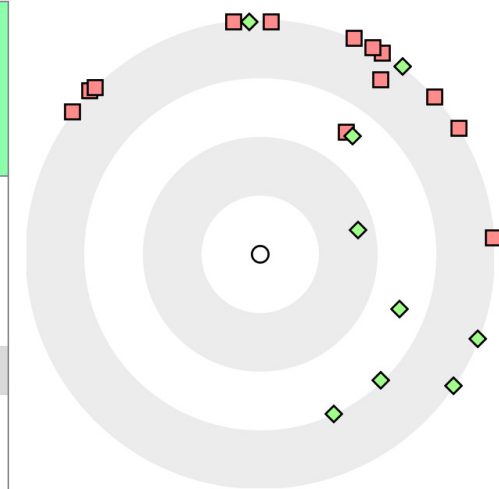
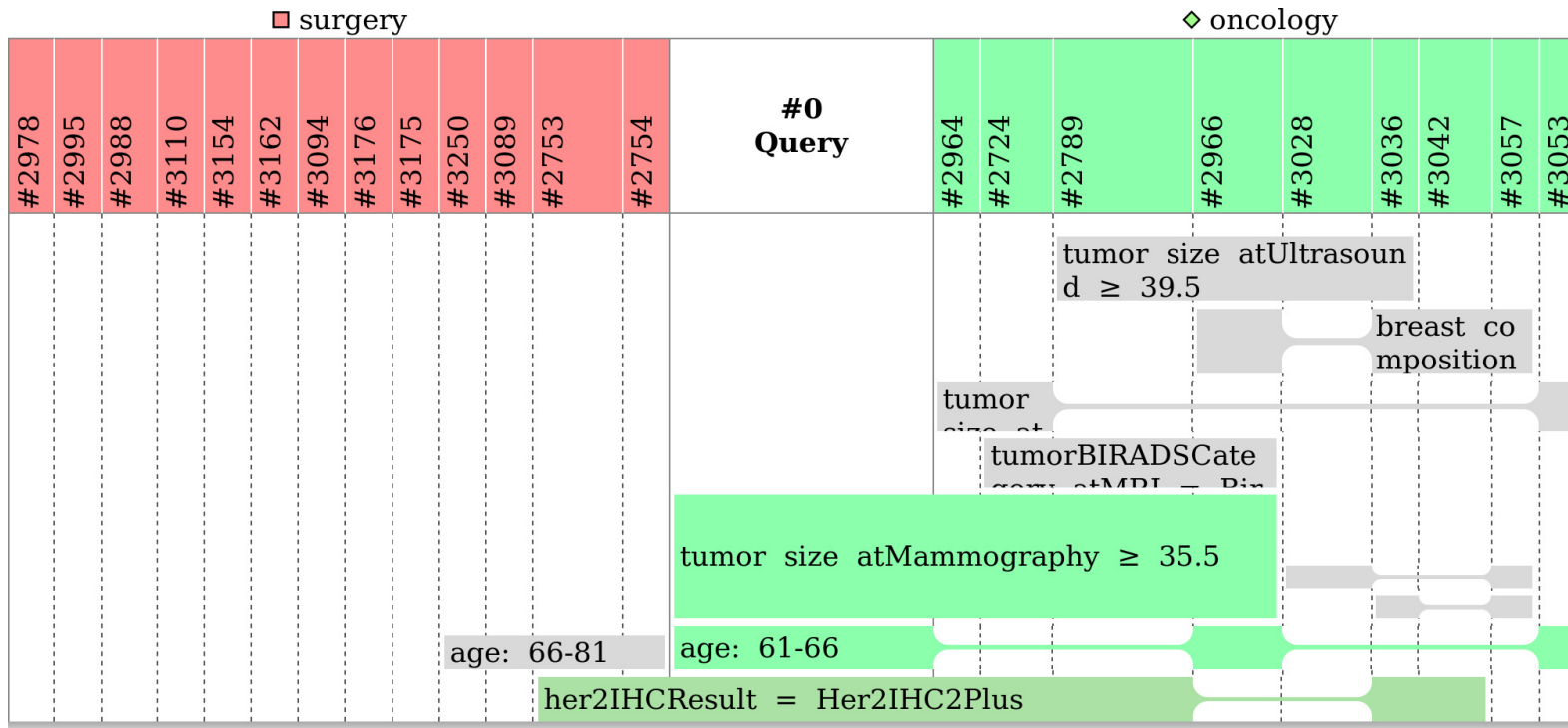
- ◆ Les boîtes donnent des arguments en faveur de l'une ou l'autre classes
- ◆ Le traitement recommandé est celui dont la couleur est majoritaire
- ◆ Le clinicien peut choisir une option différente, s'il juge les arguments non pertinents

Version hiérarchique

➤ 4 grandes classes de traitements

◆ Mais de nombreux protocoles au sein de chaque classe !

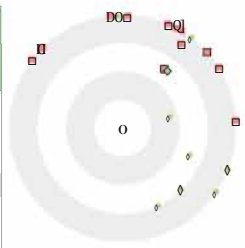
➤ Approche hiérarchique :



Compare surgery...

Compare oncology...

surgery										oncology											
#2978	#2995	#2988	#3110	#3154	#3162	#3094	#3176	#3175	#3250	#3089	#2753	#2754	#2964	#2724	#2789	#2966	#3028	#3036	#3042	#3057	#3053
										# 0 Query											
										tumor size atUltrasound: 39.0											
										breast composition											
										tumor_											
										tumorBIRADSCate											
										tumor size atMammography: 35.5											
age: 66-81										age: 61-66											
her2IHCResult = Her2IHC2Plus																					



Compare surgery...

Compare oncology...

D

Cyclophosphamide Doxorubicin
Paclitaxel Trastuzumab
standard therapy

Cyclophosphamide Docetaxel
Epirubicin

#	#	# 0 Query	#	#
age: 48-63				
tumor size atUltrasound: 37.5-48.0				
tumorBIRADSCategory atMRI - Birads5		multiple tumors atUltrasound		
tumor size atMammography: = 40.0				



Discussion et conclusion

➤ Performances équivalentes au kNN, mais explicabilité meilleure

➤ Nous avons visualisé des similarités locales

- ◆ Les similarités sont cherchées uniquement sur les cas voisins
- ◆ Elles ne sont pas forcément valables au niveau global
- ◆ Des travaux récents en IA vont dans ce sens [Zabkar]

➤ La visualisation d'ensemble paraît prometteuse pour l'IA explicable

➤ Limites :

- ◆ Nécessite une formation des médecins
- ◆ Les traitements recommandés correspondent à ceux prescrits dans le passé, mais rien ne garanti que ce soient les meilleurs

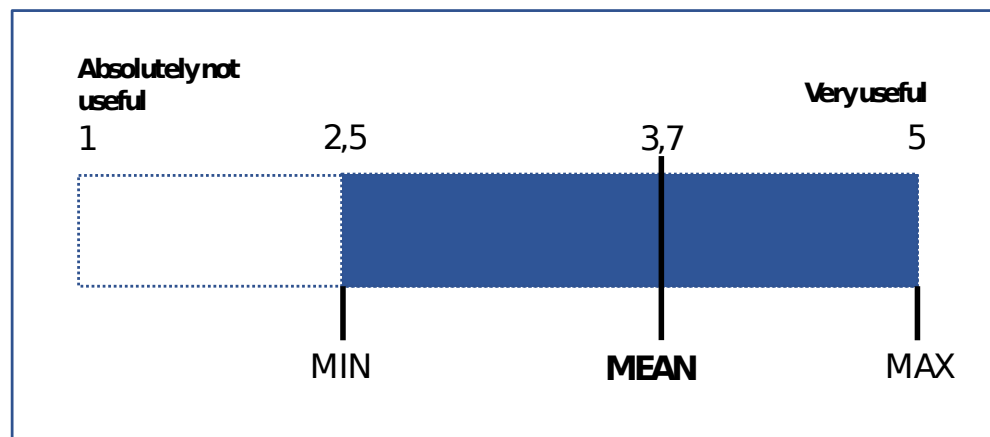
➤ Perspectives:

- ◆ Validation et évaluation cliniques
- ◆ Extension à d'autres méthodes d'IA (*deep learning, boosting, ...*)
- ◆ Adaptation à d'autres domaines

Évaluation

➤ Test auprès de 11 cliniciens

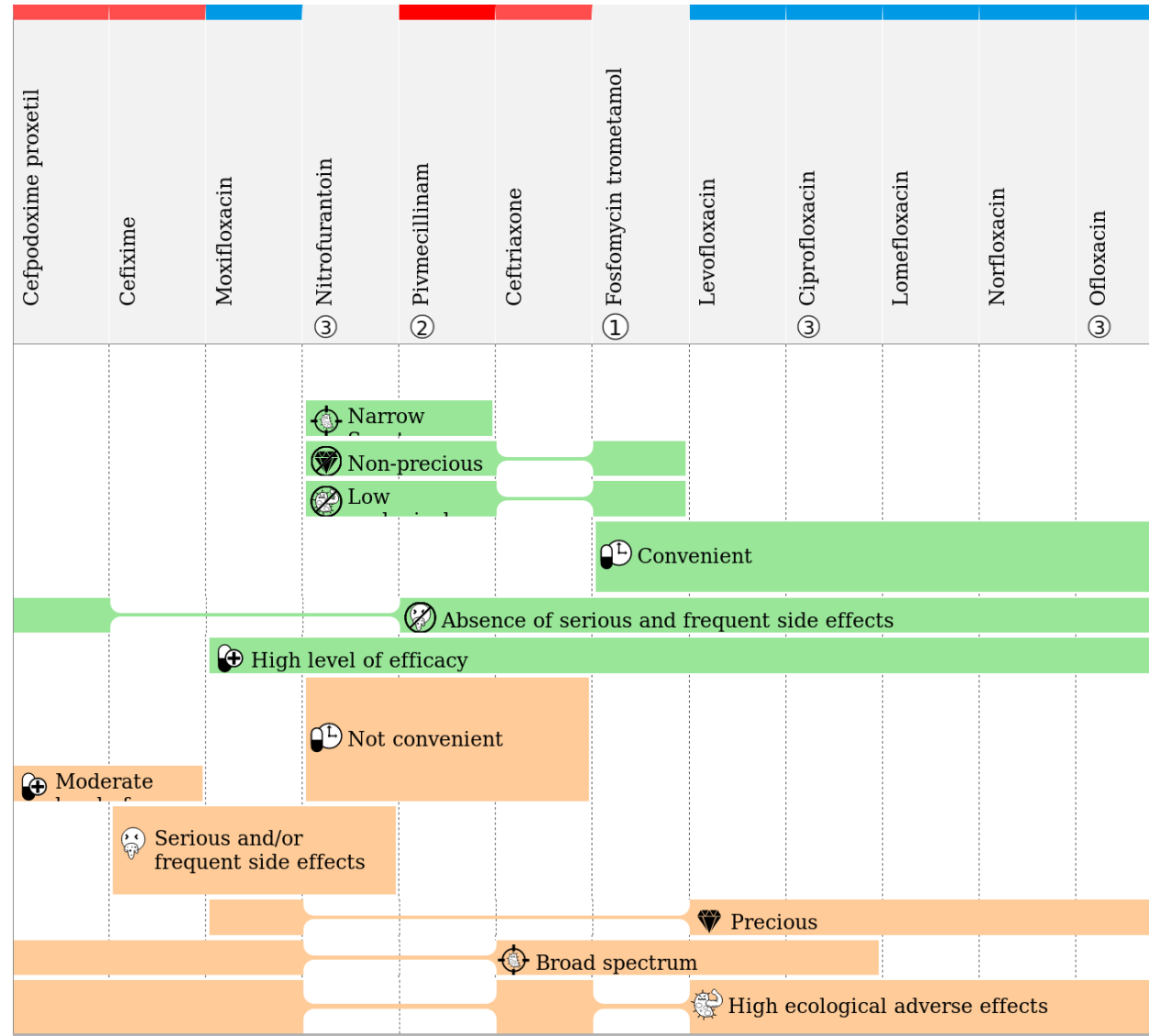
- ◆ Approche jugée très utiles pour les patients complexes, pour lesquels il n'y a pas de recommandations claires dans les guides de bonnes pratiques cliniques
- ◆ Les patients similaires retrouvés sont parfois assez éloignés du nouveau patient
 - Manque de cas dans la base
- ◆ Utilisabilité perçue :



Perspectives

➤ Aide à la prescription en antibiothérapie

◆ Visualisation des propriétés des antibiotiques



Tsopra R, Sedki K, Courtine M, Falcoff H, De Béco A, Madar R, Mechaï F, Lamy JB. Helping GPs to extrapolate guideline recommendations to patients for whom there are no explicit recommendations, through the visualization of drug properties. The example of AntibioHelp® in bacterial diseases. *Journal of the American Medical Informatics Association* 2019

References

- ◆ [AIM] Lamy JB, Sekar B, Guezennec G, Bouaud J, Séroussi B. Explainable artificial intelligence for breast cancer: a visual case-based reasoning approach. Artificial Intelligence in Medicine 2019
- ◆ [JVLC] Lamy JB, Berthelot H, Capron C, Favre M. Rainbow boxes: a new technique for overlapping set visualization and two applications in the biomedical domain. Journal of Visual Language and Computing 2017
- ◆ [iV17] Lamy JB, Tsopra R. Translating visually the reasoning of a perceptron: the weighted rainbow boxes technique and an application in antibiotherapy. International Conference Information Visualisation (iV) 2017
- ◆ [Springer] Lamy JB. Artificial Feeding Birds (AFB): a new metaheuristic inspired by the behavior of pigeons. In : Advances in nature-inspired computing and applications 2019
- ◆ [Zabkar] Zabkar J, Bratko I, Demsar J. Extracting qualitative relations from categorical data. Artificial Intelligence 2016
- ◆ [Klawonn] Klawonn F, Lechner W, Grigull L. Case-centred multidimensional scaling for classification visualisation in medical diagnosis. International Conference on Health Information Science 2013

Questions et discussion

