



Université de La Rochelle
Pôle Sciences & technologies



Laboratoire L3I
Informatique, Image et Interaction
EA 2118

Structure de treillis

Contributions structurelles et algorithmiques
Quelques usages pour des données images

Habilitation à Diriger les Recherches
14 Juin 2011

Présentée par **Karell Bertet**

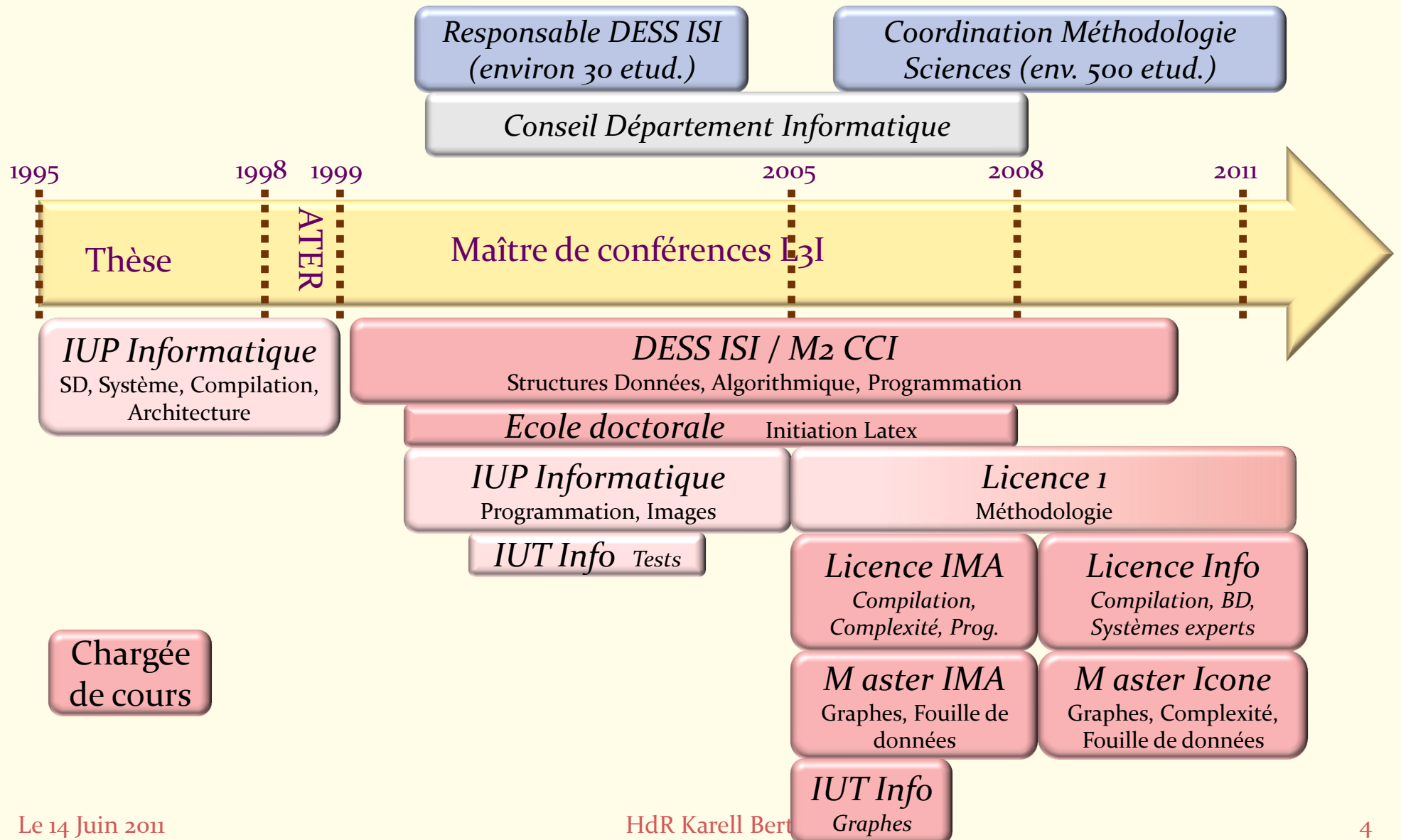
Eléments de CV

1995-2011

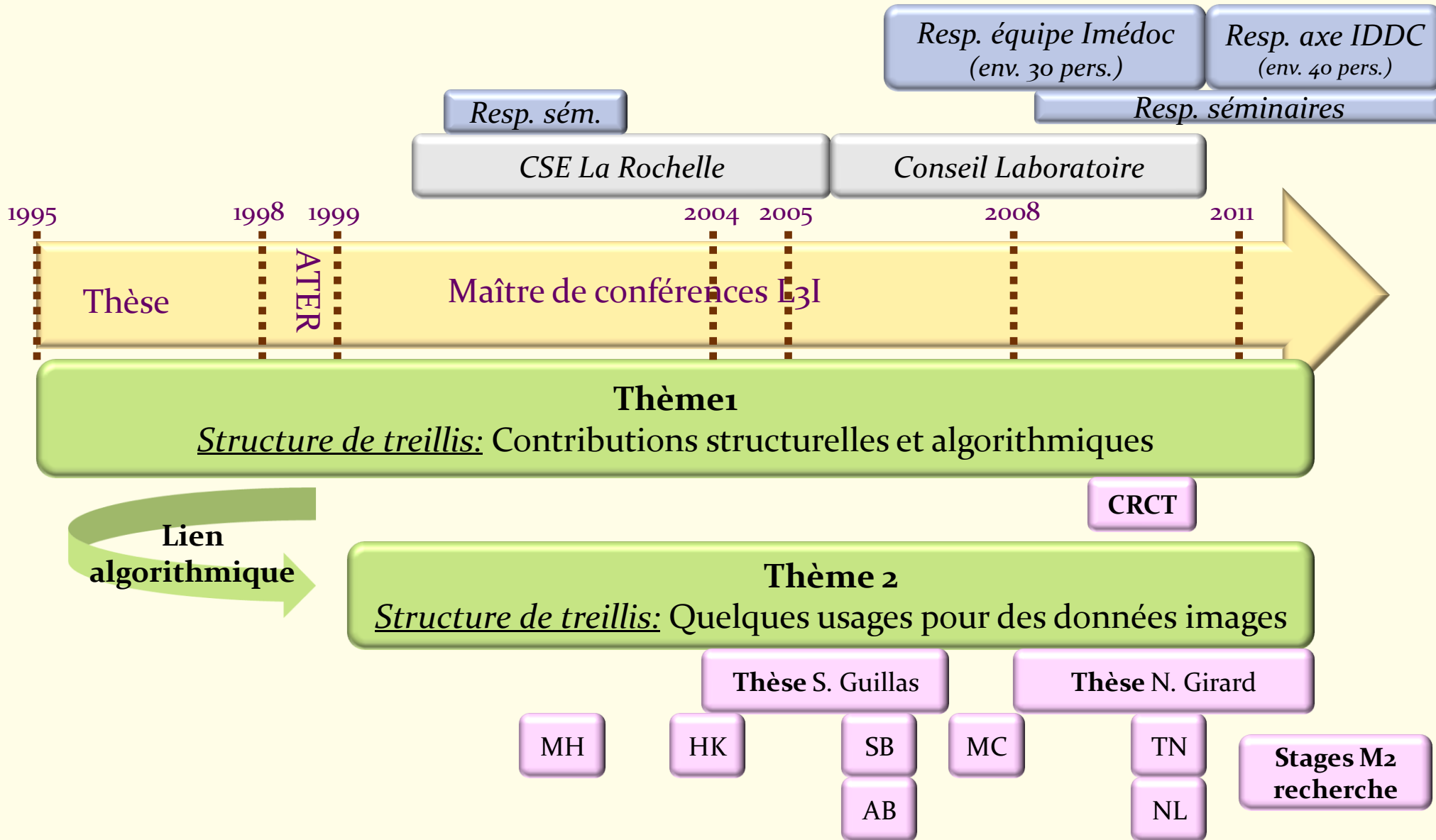
Quelques dates



Enseignement



Recherche



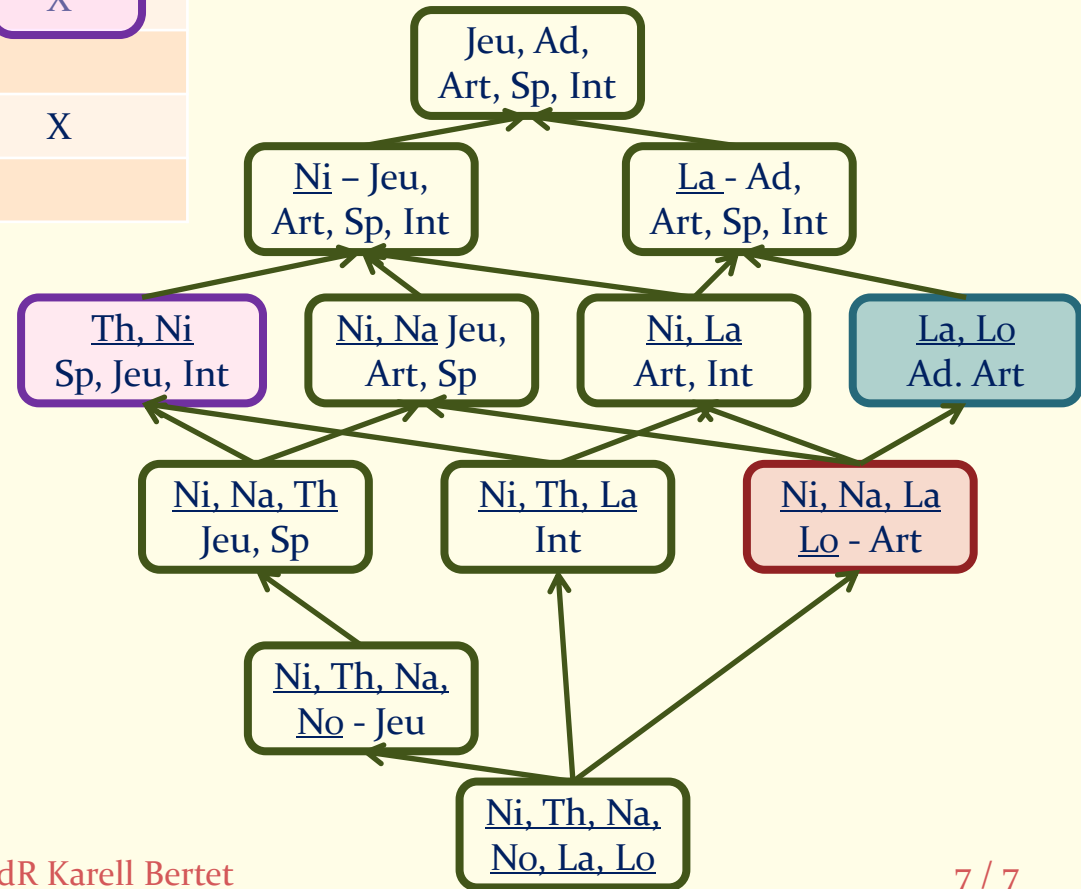
Thème 1

Structure de treillis:

Contributions structurelles et algorithmiques

Un exemple introductif

	Tranche d'âge		Activités		
	<u>J</u> eune	<u>A</u> dulte	<u>A</u> rtistique	<u>S</u> portive	<u>I</u> nternet
<u>N</u> oam	X				
<u>T</u> héo	X			X	X
<u>N</u> ina	X		X	X	X
<u>N</u> ais	X		X	X	
<u>L</u> aurent		X	X		X
<u>L</u> ouise		X	X		



Problématique

Identification d'un **jeu algorithmique** adapté et efficace pour manipuler les objets de la théorie des treillis

- ⇒ Etude des liens structurels entre ces objets
- ⇒ Décloisonnement scientifique

		attributs S			
		a	b	c	d
objets O	1	x		x	
	2			x	x
	3		x		x

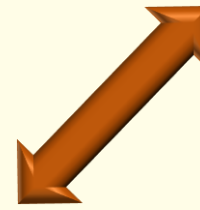
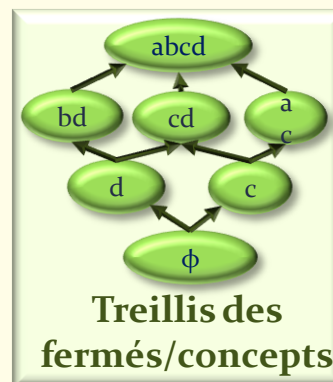
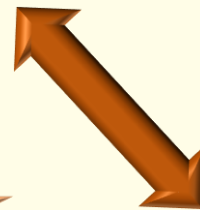
Table binaire
contexte

b	→	d
bc	→	a
a	→	c
ad	→	b

Règles
d'implication



Informatique:
Des données au
treillis



Math. discrètes:
Du treillis aux
représentations

Des données au treillis des concepts

		attributs S			
		a	b	c	d
objets O	1	x		x	
	2			x	x
	3		x		x

Table binaire/contexte
(O,S,(α,β))

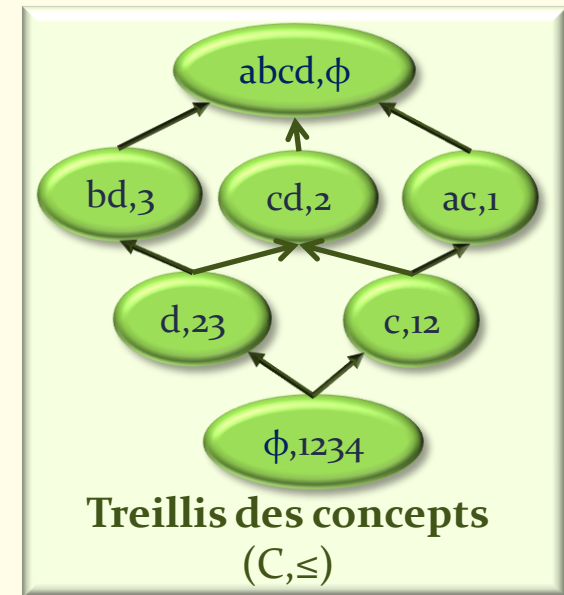


Table binaire / contexte (O,S,(α, β))

- $\alpha : P(O) \rightarrow P(S)$

α associe à des objets leurs attributs communs

- $\beta : P(S) \rightarrow P(O)$

β associe à des attributs les objets qui les partagent

Treillis des concepts (C, \leq) [GW99]

- **C**: ensemble de concepts

Un concept est une paire (A,B) avec $A \subseteq O$, $B \subseteq S$, $B = \alpha(A)$ et $A = \beta(B)$

- \leq : relation binaire de « spécialisation / généralisation » entre concepts
 $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow B_1 \subseteq B_2 \Leftrightarrow A_1 \supseteq A_2$

Des données au treillis de Galois

		attributs S			
		a	b	c	d
objets O	1	x		x	
	2			x	x
	3		x		x

Table binaire/contexte
(O,S,(α , β))

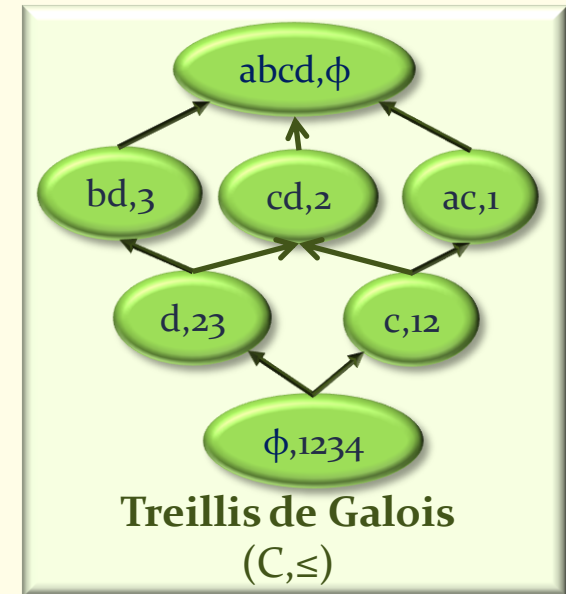


Table binaire / contexte (O,S,(α , β))

- (α , β) forme une **connexion de Galois** entre $P(O)$ et $P(S)$

Treillis de Galois (C, \leq) [BM70]

- **C**: ensemble de paires de fermés sur $P(O) \times P(S)$
(A,B) \in C si $B = (\beta \circ \alpha)(A)$ et $A = \beta(B)$

Treillis de Galois et treillis des concepts

Treillis des concepts [GW99]:

- Définition ensembliste: Objets \leftrightarrow ensembles d'attributs binaires

Treillis de Galois [BM70]:

- Connexion de Galois: Objets \leftrightarrow attributs binaires
 \Rightarrow Objets \leftrightarrow autres descriptions

• Exemples :

- Objets \leftrightarrow intervalles, histogrammes [Pol98]
- Objets \leftrightarrow formules logiques (ALC) [FR04]
- Objets \leftrightarrow graphes [GS01]
- Objets \leftrightarrow « patrons » [GK01]

Des données au treillis des fermés

		attributs S			
		a	b	c	d
objets O	1	x		x	
	2			x	x
	3		x		x

Table binaire/contexte
(O,S,(α,β))

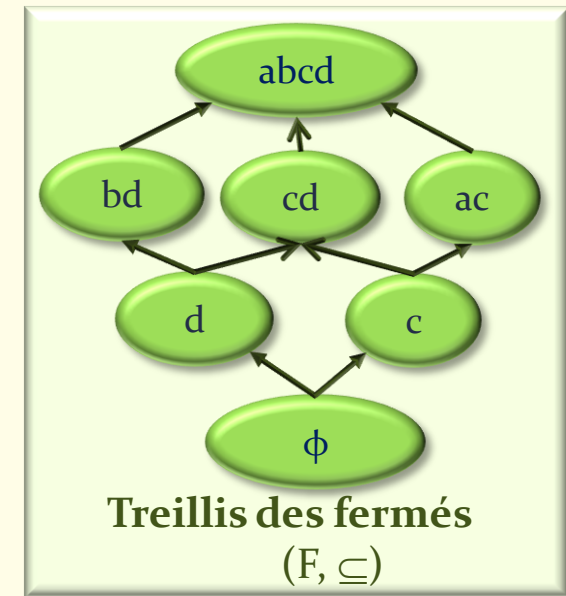
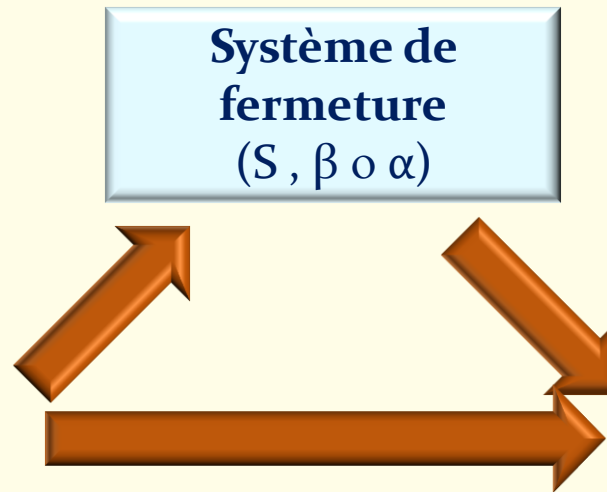


Table binaire / contexte (O,S,(α,β))

Décrite par deux opérateurs de fermeture:

- ($\beta \circ \alpha$) opérateur de fermeture sur S
- ($\alpha \circ \beta$) opérateur de fermeture sur O

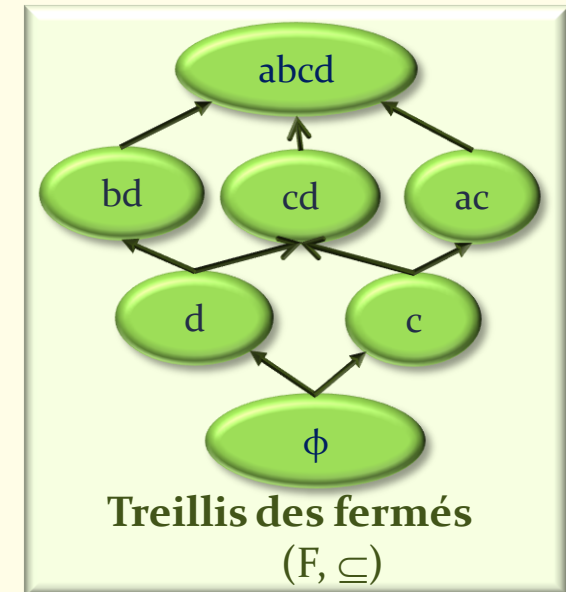
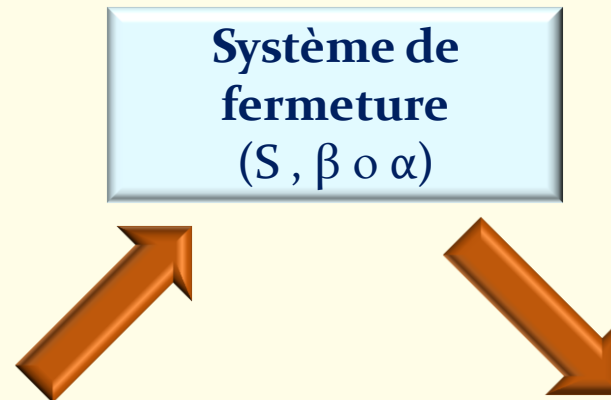
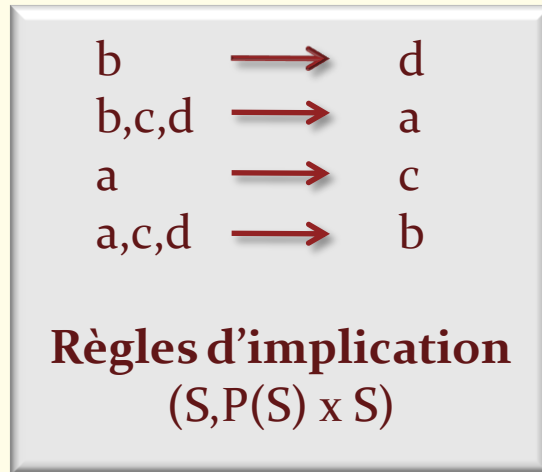
Treillis des fermés (F, \subseteq) [MC03]

• **F**: ensemble de fermés sur S

$X \subseteq S$ est un fermé si $X = (\beta \circ \alpha)(X)$

F est une famille de Moore sur S

Des règles au treillis des fermés



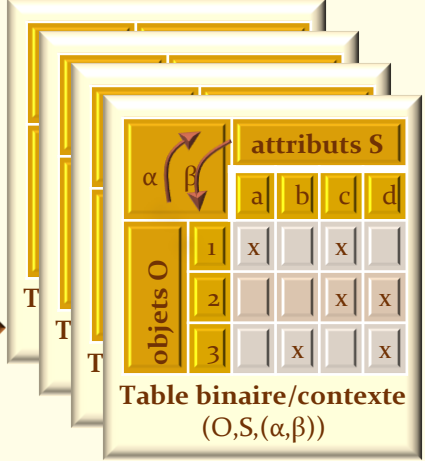
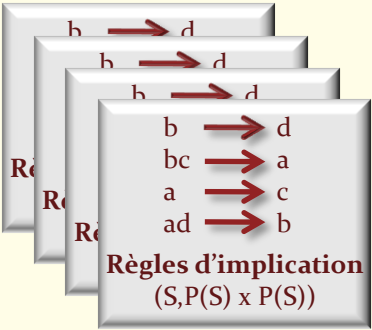
Règles d'implication ($S, P(S) \times S$):

- **Implication:** notée $b, c \rightarrow a$,
« *b et c impliquent a* »
- **Opérateur de fermeture δ sur $P(S)$:**
 $\rho(X)$ = application des règles pour X

Treillis des fermés (F, \subseteq) [MCo3]

- **F:** ensemble de fermés sur S
- **Fermés:** $X \subseteq S$ tq $X = \rho(X)$

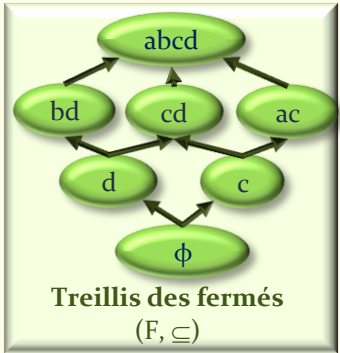
Positionnement



Système de fermeture
(S, ρ)



Informatique:
Des données au treillis



Math. discrètes:
Du treillis aux représentations

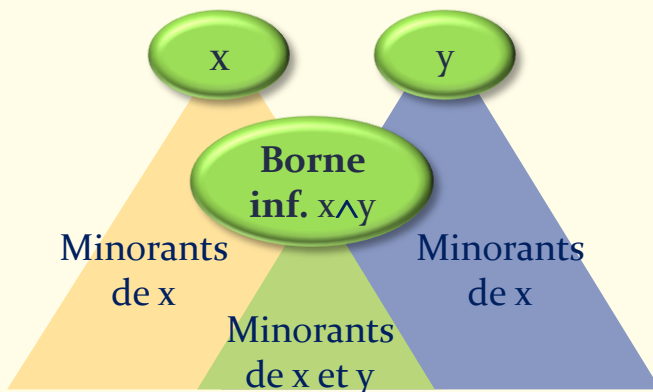
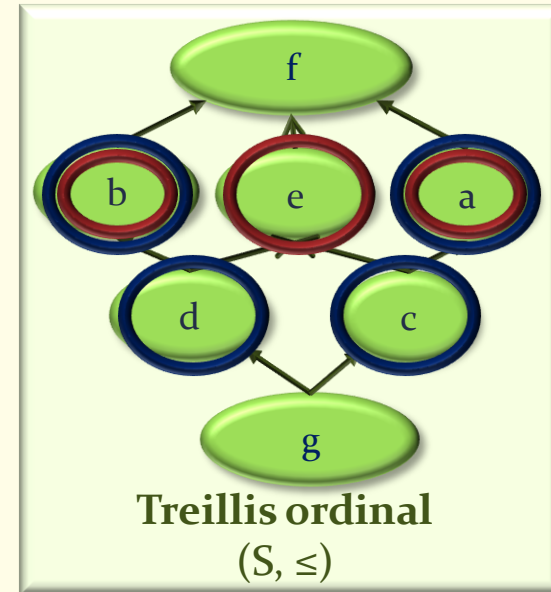
Treillis algébrique et ordinal

Treillis algébrique (S, \wedge, \vee) [Ded1903] :

- S : ensemble d'éléments
- \wedge, \vee : opérateurs de bornes supérieure et inférieure

Treillis ordinal (S, \leq) [Bir40]:

- S : ensemble d'éléments
- \leq : relation d'ordre tq toute paire (x, y) possède:
 - une borne supérieure
 - une borne inférieure



Élément irréductible: élément qui n'est la borne supérieure ou inférieure d'aucune partie ne le contenant pas

- **sup-irréductibles** J
- **inf-irréductible** M

\Rightarrow Les ensembles J et M décrivent la structure du treillis

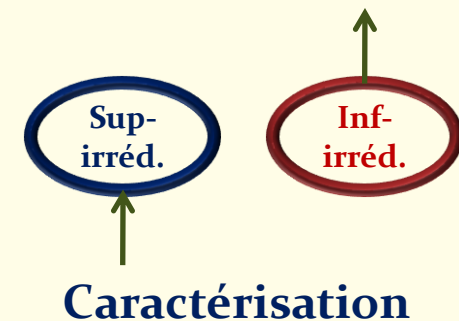
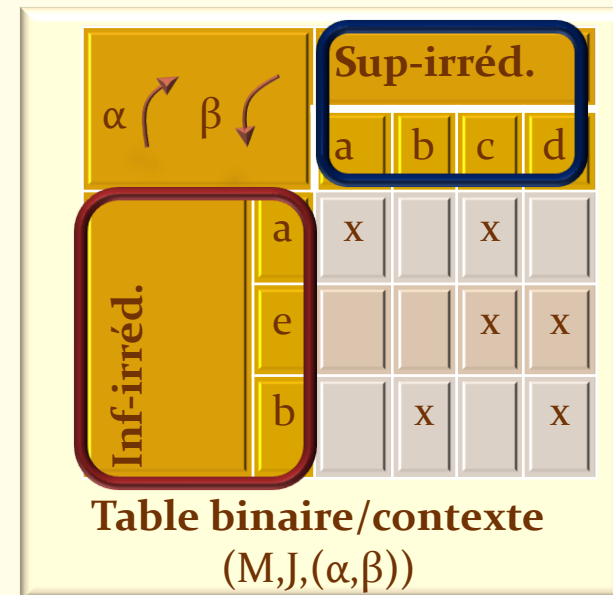
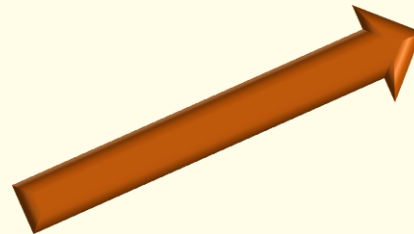
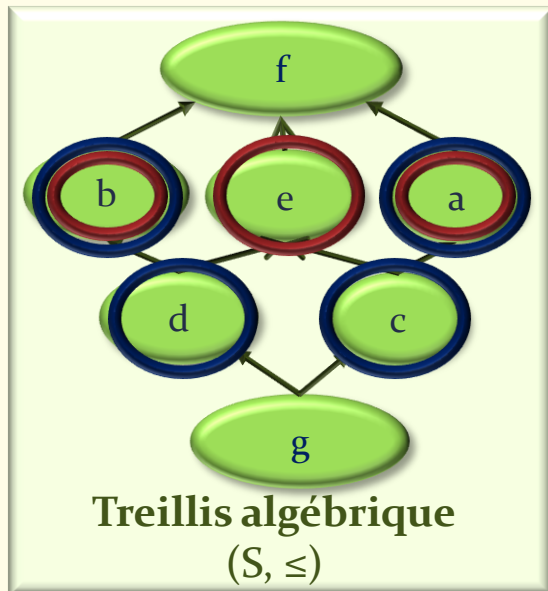


Table d'un treillis

Table binaire d'un treillis:

- Sup-irréductibles J en colonne
- Inf-irréductibles M en ligne
- $\beta(j) = \{ m \in M : j \leq m \}$



Théorème [BM70]

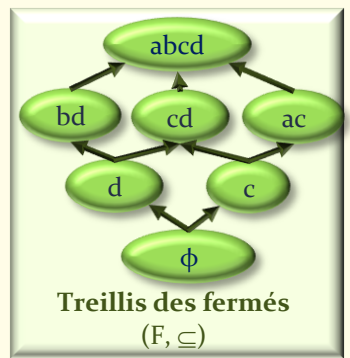
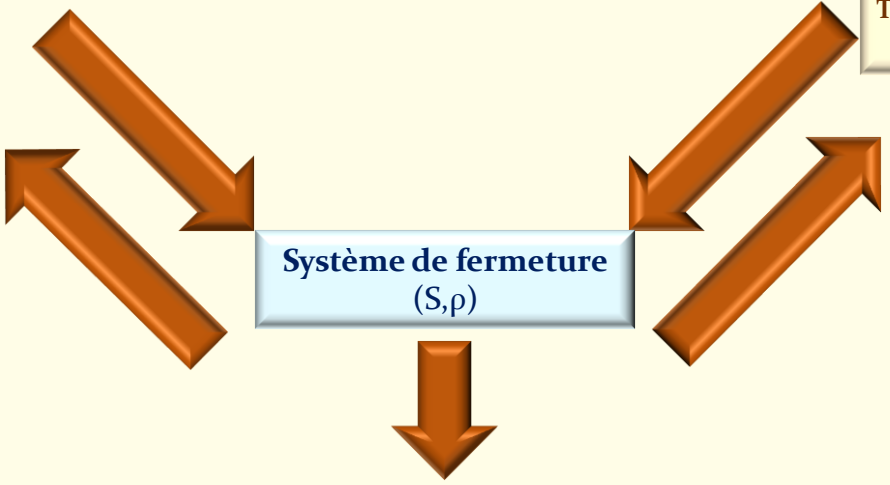
Tout treillis est isomorphe au *treillis des concepts* de sa table binaire

\Rightarrow Bijection

Treillis \leftrightarrow Table binaire *réduite*

Positionnement

- Base canonique [GD82]
- Base canonique et « iteration-free » [Wild94]
- Base faible d'implications [RW95]
- Dépendances fonctionnelles minimales [Mai83]
- Base d'implications propres [TB02]



**Math. discrètes:
Du treillis aux représentations**

Contributions

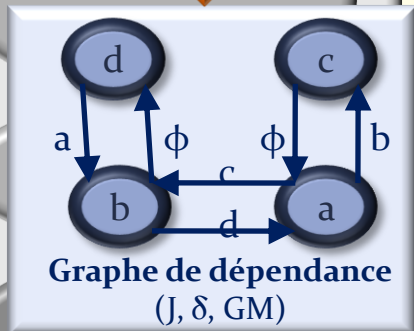


Base canonique

Base canonique directe [Bertet et al:TCS 10]

d'implications [RW9]

Dépendances fonctionnelles

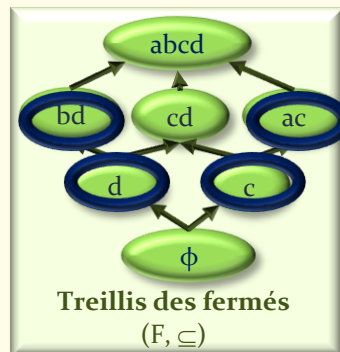


dépendance [Bertet 98]

Base directe-optimale [Bertet et al. 04]

Unicité [Bertet et al.10]

Système de fermeture (S, ρ)

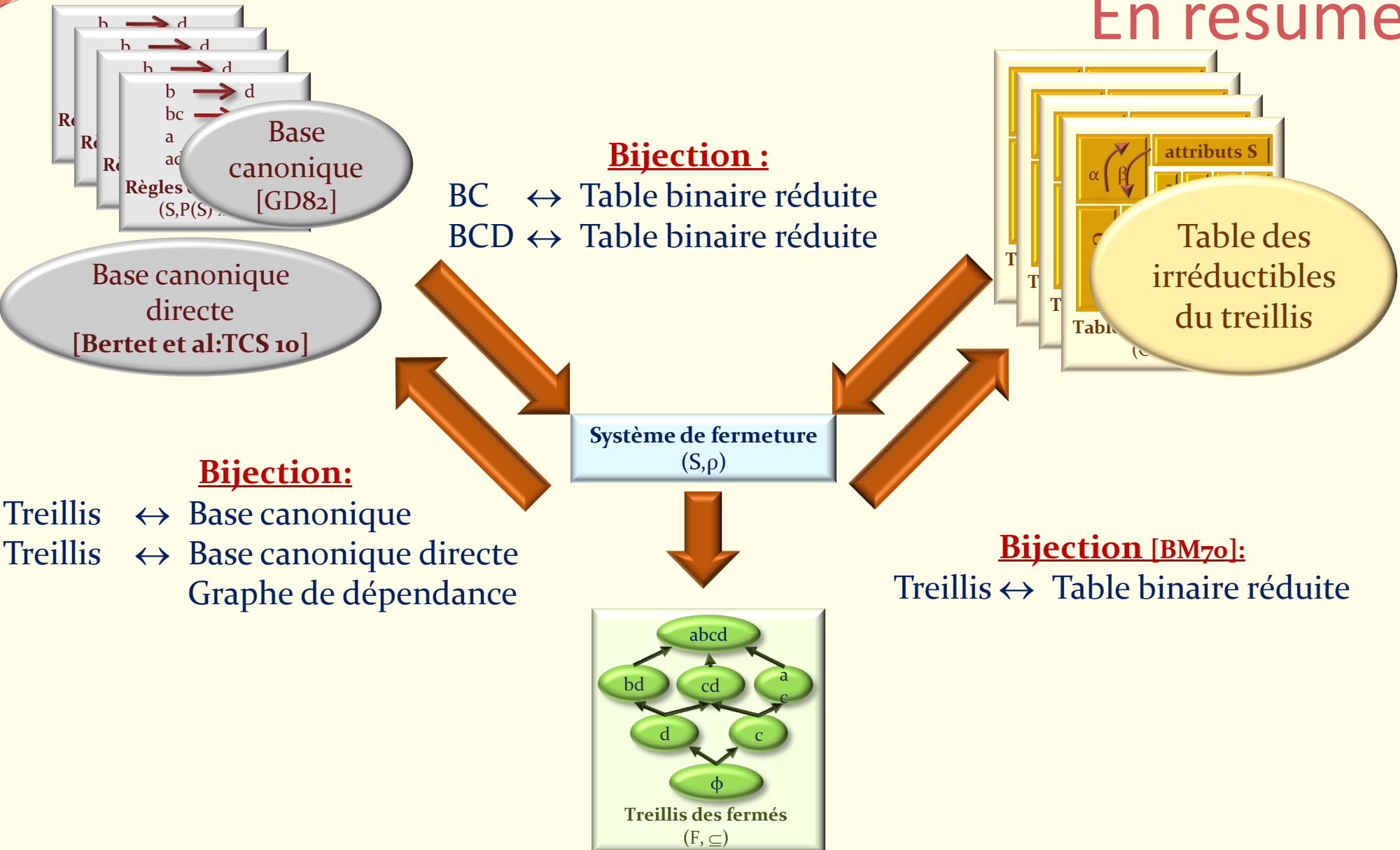


Math. discrètes:
Du treillis aux représentations

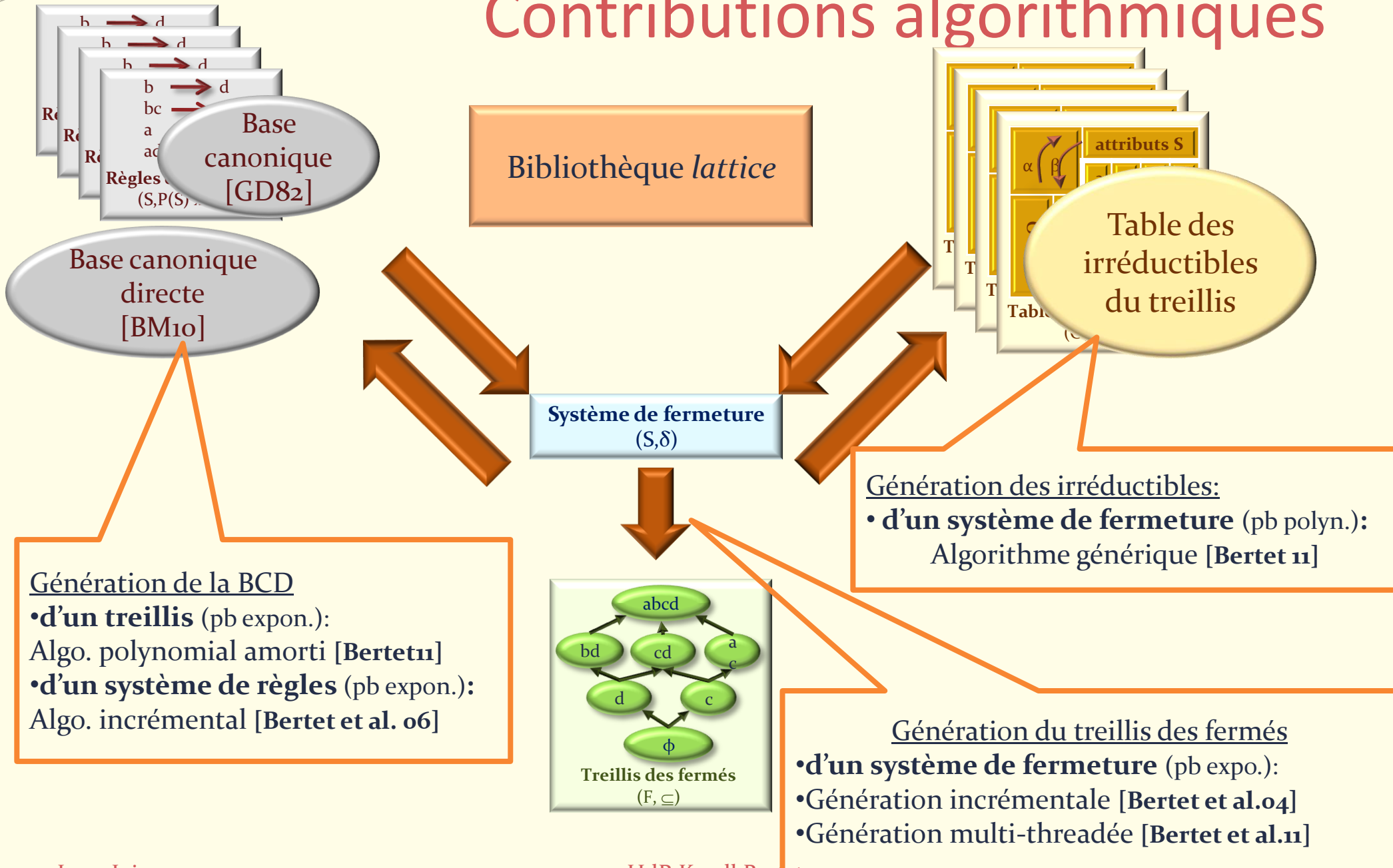
Contributions structurelles:

- Autour de la **définition ordinale**: [Bertet et al:Order02] [Bertet et al.99] [Bertet 98] [Bertet et al.97]
- Autour de l'**utilisation algorithmique** d'un treillis: [Bertet et al:Algebra01] [Bertet et al:TCS00]

En résumé



Contributions algorithmiques



Conclusion

• En théorie des treillis, différentes terminologies pour différents usages:

- **Treillis algébrique et ordinal** [Dedo7, Bir40]
 - Irréductibles et classes de treillis
- **Treillis de Galois** [BM70]
 - Connexion de Galois
- **Treillis des fermés** [MCo3]
 - Bases de règles
- **Treillis des concepts** [GW99]:
 - Analyse Formelle des Concepts

Importance du lien entre les objets issus de la théorie des treillis:

⇒ Retombées **algorithmiques** pour une utilisation efficace

⇒ **Décloisonnement communautaire** mathématiques discrètes/informatique

Thème 2

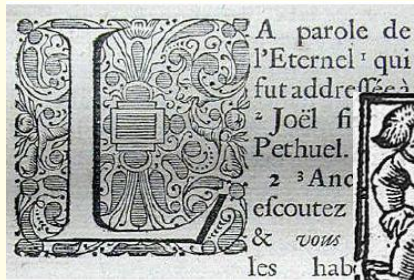
Structure de treillis:

Quelques usages pour des données images

Images de documents

- Images issues d'un **domaine particulier**

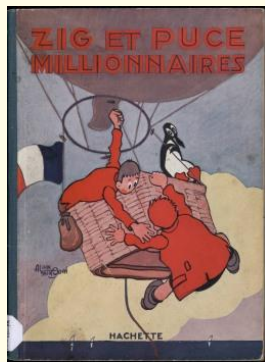
- Livres anciens (lettrines)



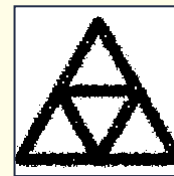
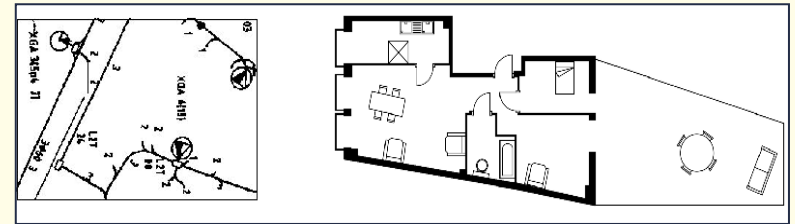
ANR
Madonne
(04-06)

- Bandes dessinées

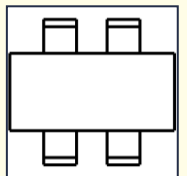
Action
CPER
2011



- Documents techniques (symboles)



ANR
Navidomass
(06-09)



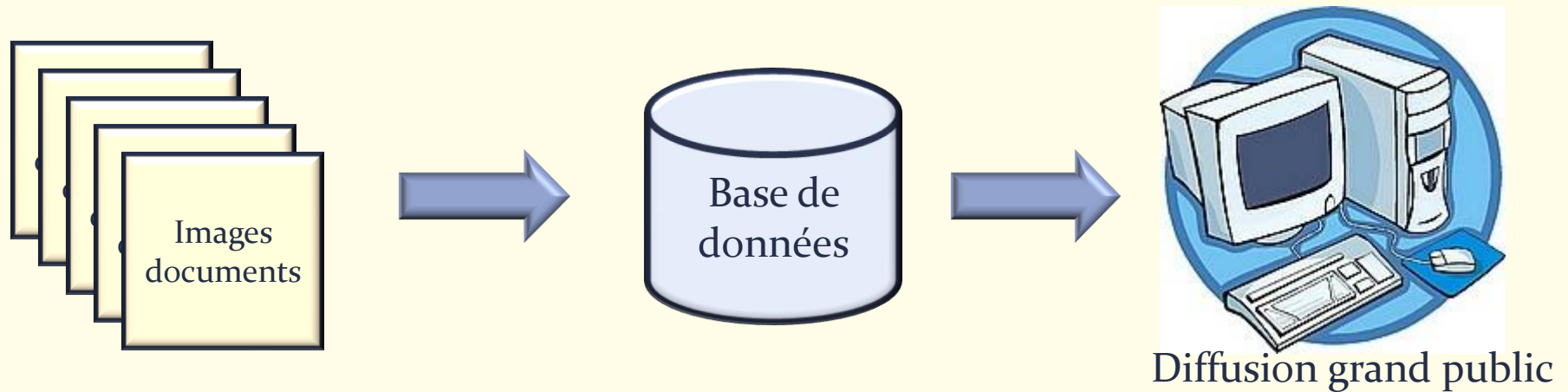
- Particularités:

- Images fortement structurées (information spatiale)
- Images à bases de traits
- Images détériorées (numérisation, dégradations, ...)

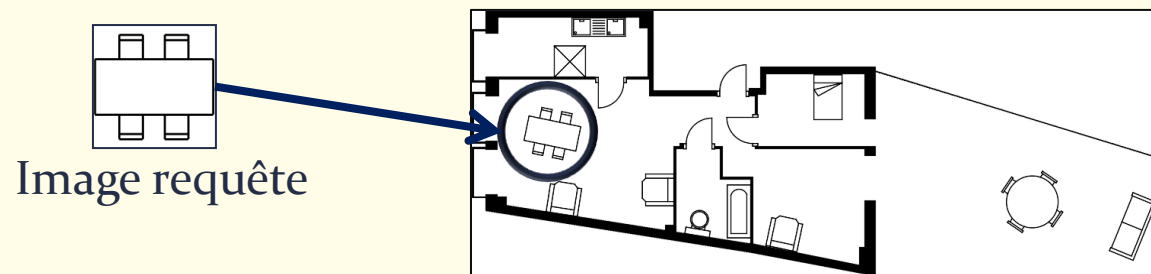
⇒ Méthodes spécifiques d'extraction de **signatures**

Problématiques

- Valorisation du patrimoine



- Reconnaissance en contexte (CBIR)



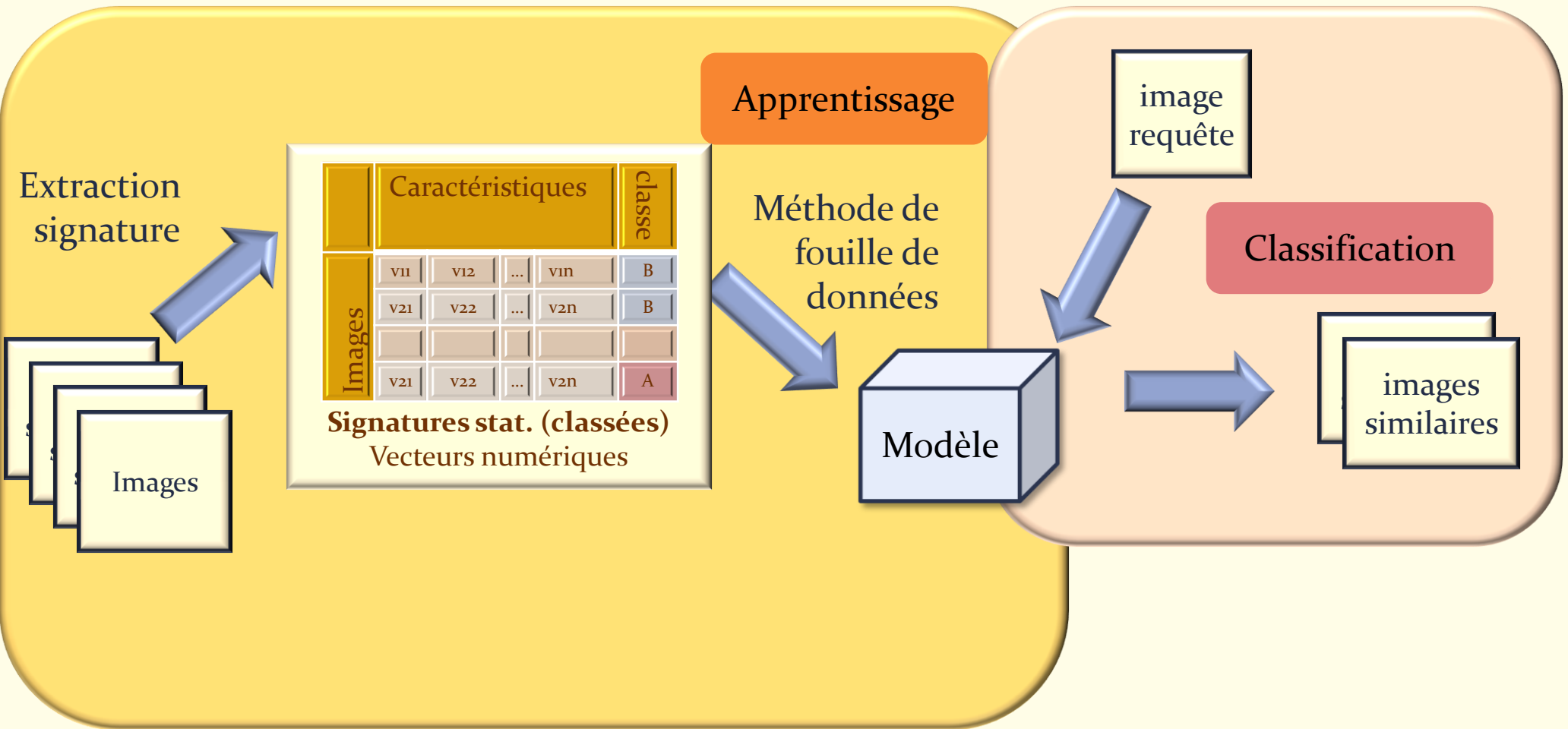
Problématiques

- Images de documents & **Fouille de données**
 - Objectif: Classification / indexation / sélection d'attributs
 - Lisibilité des méthodes (pouvoir d'explication)
 - Intégration d'information spatiale (CBIR)
- Images de documents & **Représentation des connaissances**
 - Objectif: Recherche d'information
 - Intégration de données:
 - numériques (signatures)
 - symboliques (sémantique du domaine, information spatiale)

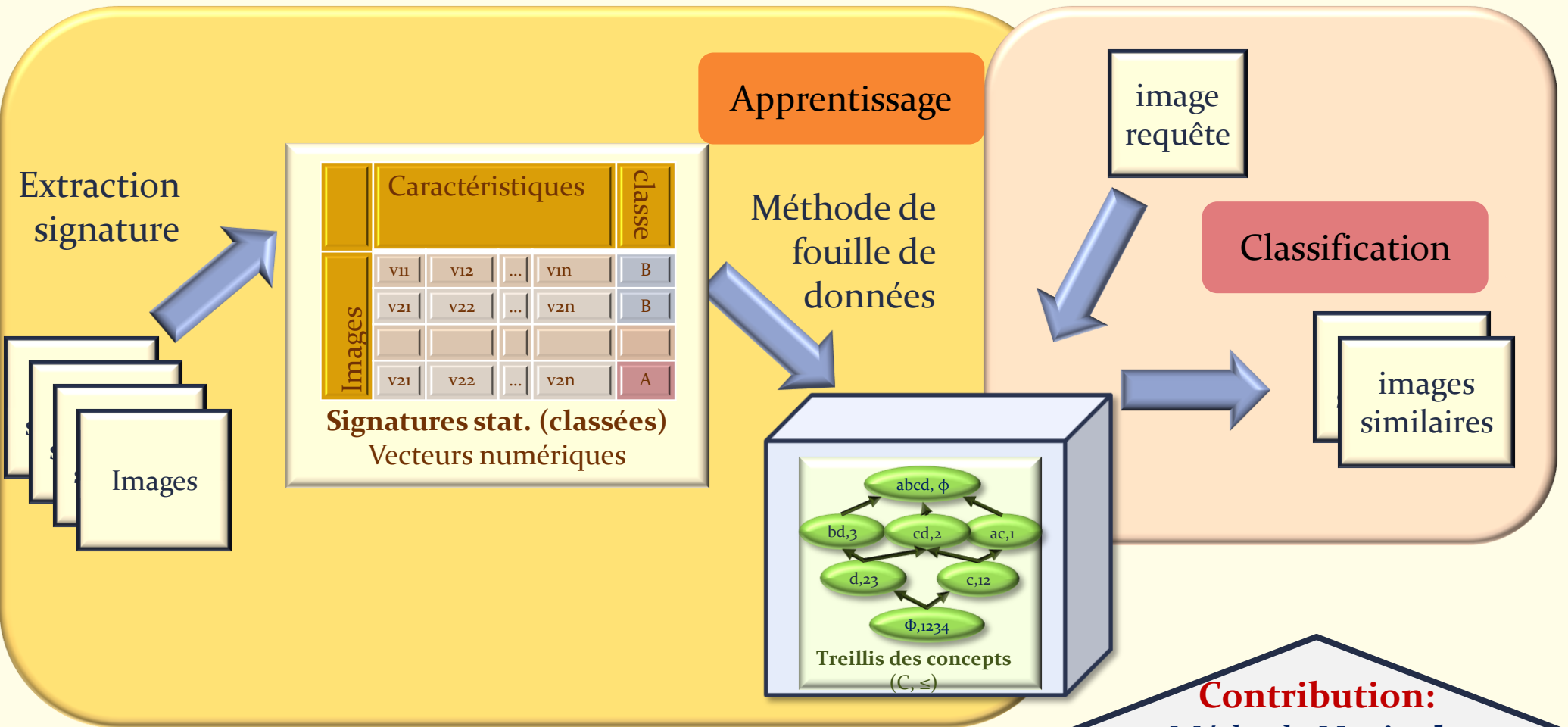
Intérêt des **méthodes de l'AFC** (Analyse Formelle des Concepts):

- Objectif: Classification, Représentation des connaissances
- Lisibilité (explication par des règles)
- Méthodes symboliques (données numériques et symboliques)

Classification et signature statistique



Classification et signature statistique



Contribution:
Méthode Navigala
de classification par
navigation dans un
treillis
[Bertet et al:IJPRAI 10]

Thèse S. Guillas [04-07], Thèse N. Girard en cours, 4 stages M2
1 revue inter, 2 revues nat, 7 conférences, 3 lncs, 1 logiciel

Contribution: méthode Navigala

Caractéristiques					classe
Images	v11	v12	...	v1n	B
	v21	v22	...	v2n	B
	v21	v22	...	v2n	A

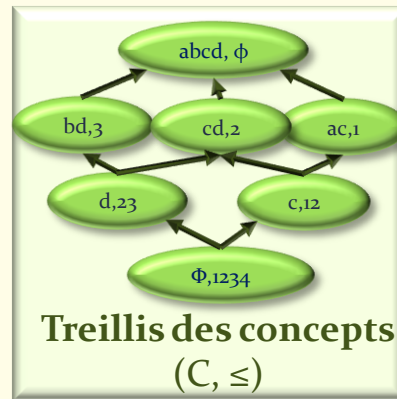
Signatures stat. (classées)
Vecteurs numériques

Discrétisation

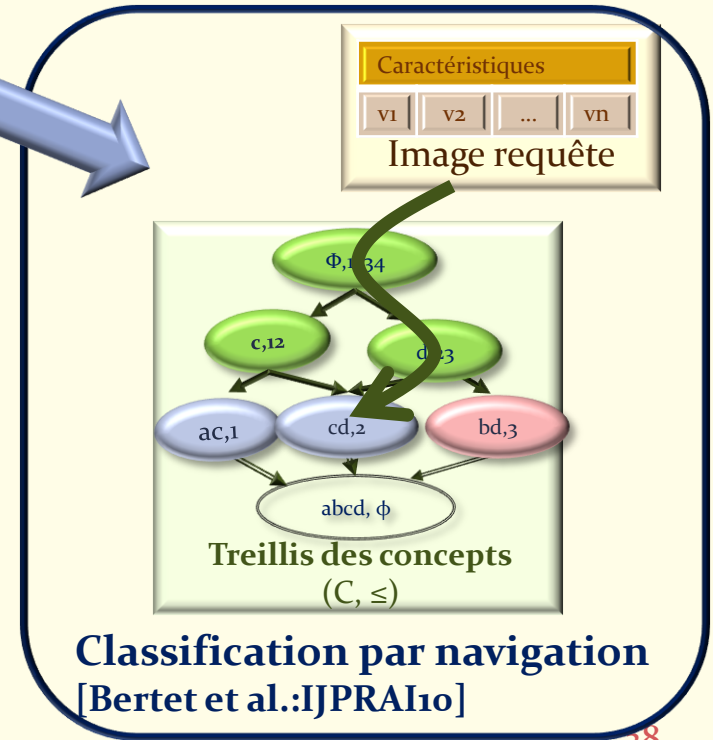
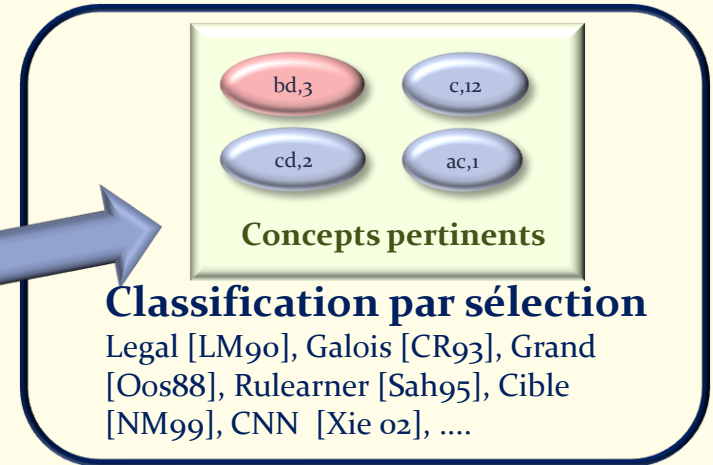
Génération du treillis

		attributs S				Classe
		c1	cn			
		a	d	b	c	
Images O	1	X			X	B
	2		X		X	B
	3		X	X		A

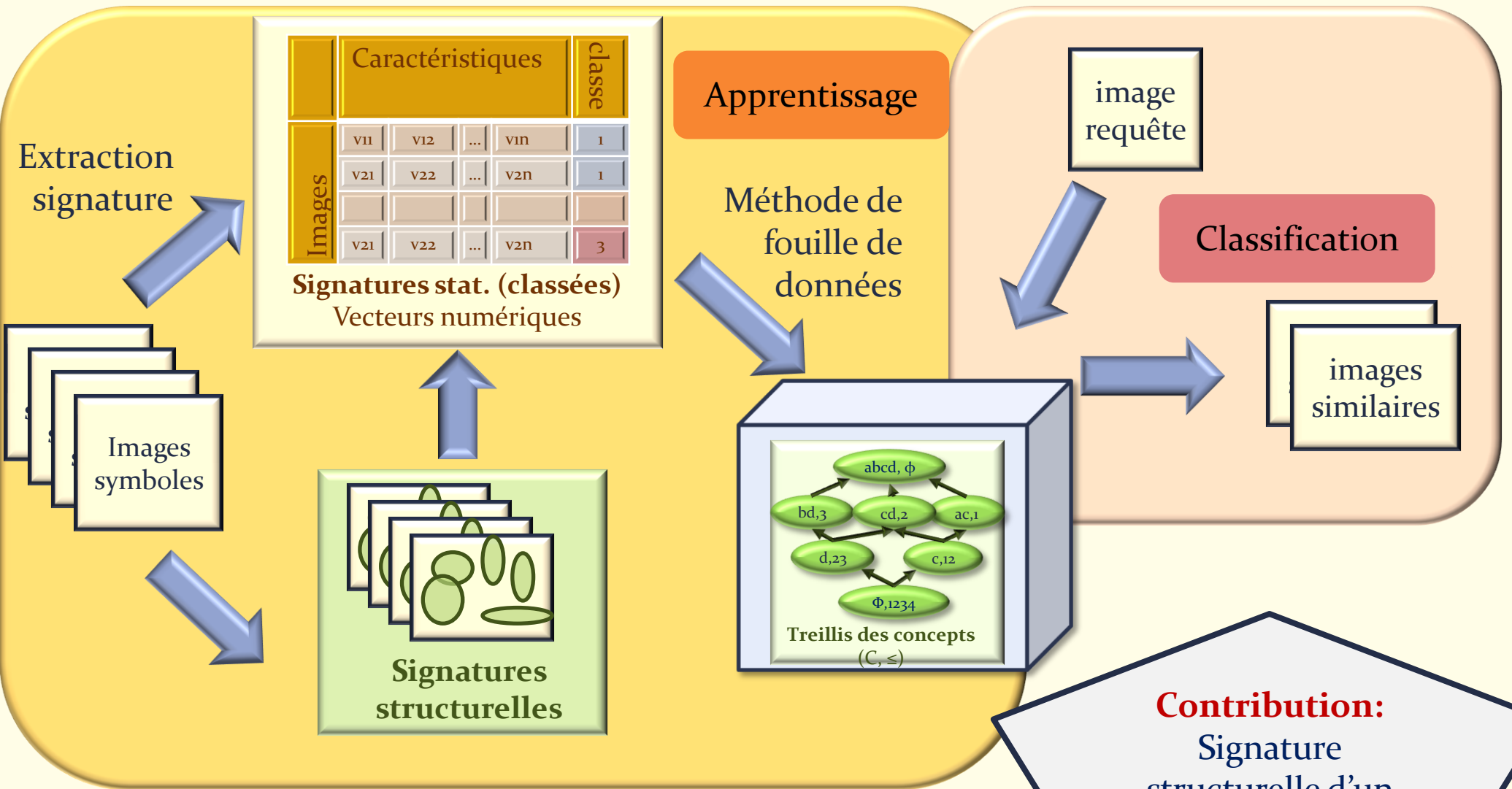
Table binaire/contexte
(O,S,(α,β))



- **Mécanisme de navigation**
(treillis des fermés)
- **Génération à la demande**
(génération des concepts successeurs)
- **Discrétisation locale**
(irréductibles)
- **Treillis dichotomiques**
(co-atomisticité)



Classification et signature structurelle



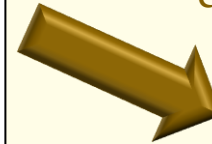
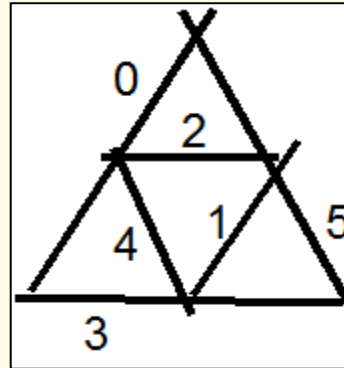
Contribution:
Signature structurelle d'un symbole
[Bertet et al.:SMC11]

*Thèse S. Guillas [04-07], 2 stages M2
1 revues intern., 1 revue nat., 2 conférences, 1 lncs*

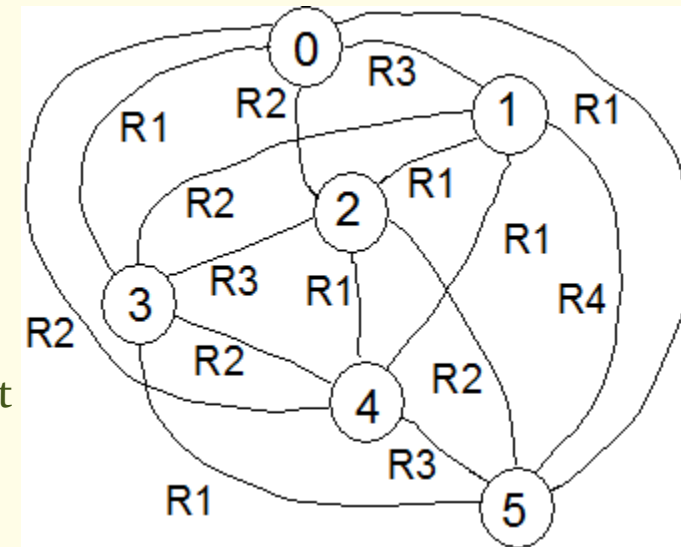
Contribution: signature structurelle



Squelettisation +
Extraction de segments
maximaux par transformée de
Hough [Hou59]



Graphe topologique
complet (5 relations
topologiques)



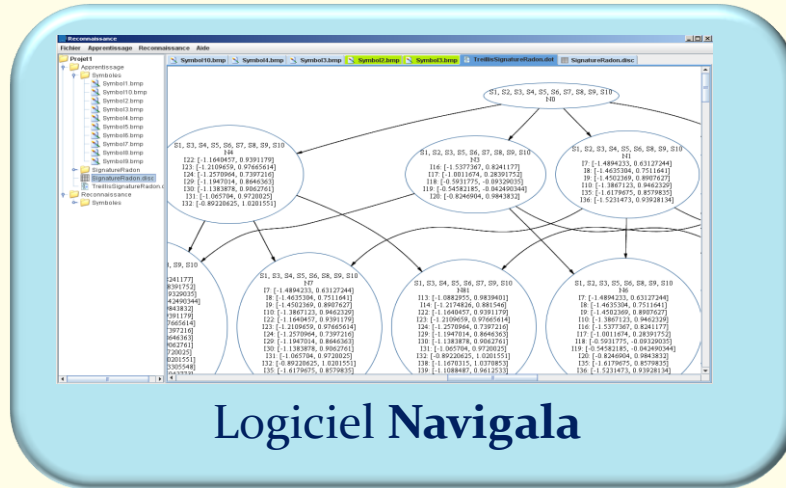
Extraction de chemins /
formes (adaptation du produit
de matrices)

Chemins	$R_3R_1R_3$	R_2R_3	...
Occurrences	6	10	...

Signature

- Pas de connaissance a priori
- Possibilités d'extension à d'autres primitives
- Invariance rotation, translation, changement d'échelle

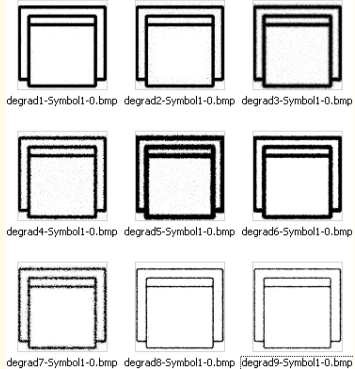
Expérimentations



Logiciel Navigala

Bases:

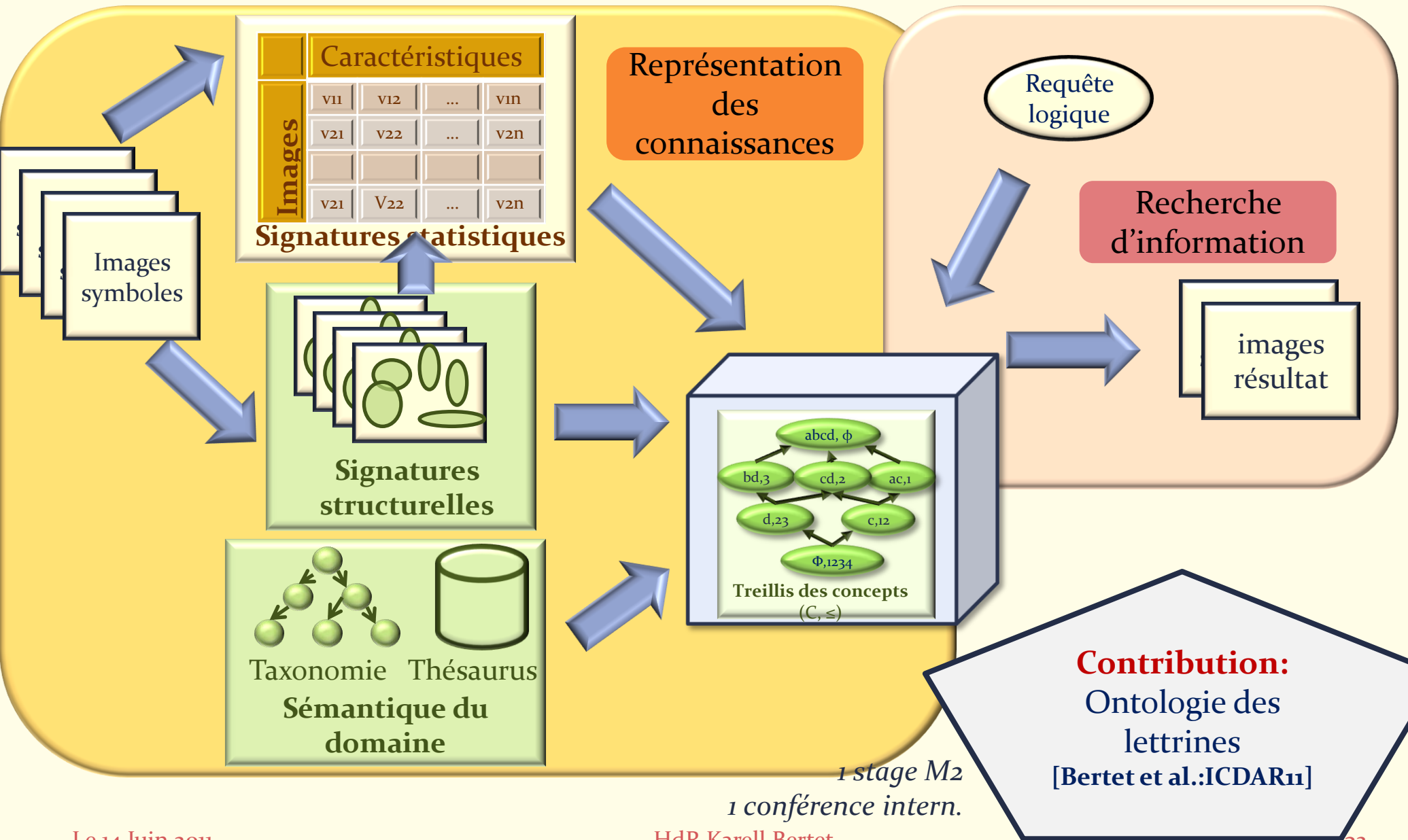
•GREC [Greco3]



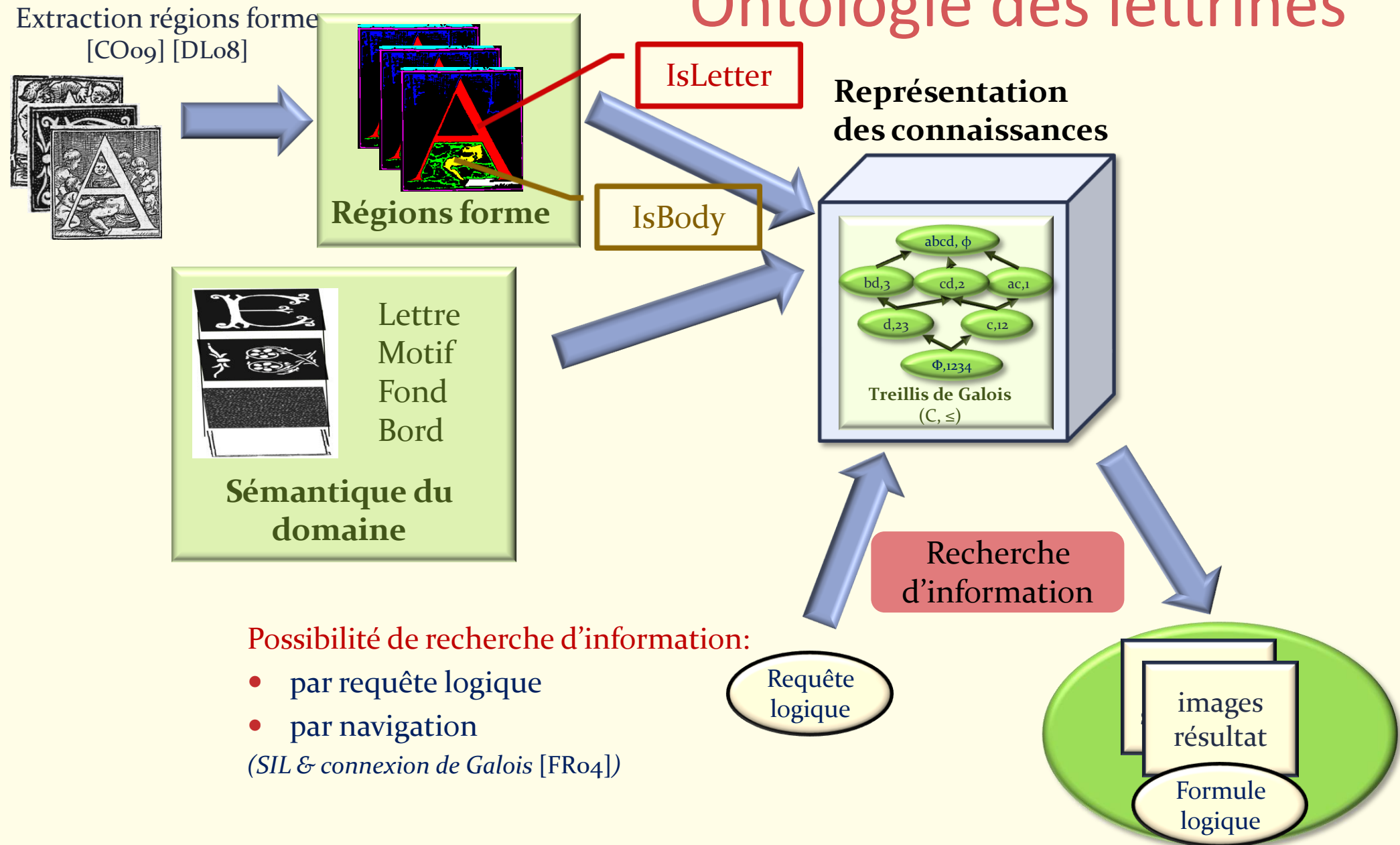
•UCI repository [UCI₁₀]

Méthodes	Navigala (globale)		Navigala (locale)		SVM	AD (Chaid)
	Nb classes	Taux reco.	Nb concepts	Taux reco.	Taux reco.	Taux reco.
Bases						
Greco (Radon)	10	81,42 %	2192	90,43 %	92	98,13 %
Greco (structurelle)	19	73,87 %	4308	74,60 %	3851	53,08%
GLASS	6	73,18 %	2074	72,60 %	2128	68,65 %
IRIS	3	93,33 %	195	96,67 %	42	95,33 %
Breast cancer	2	90,05 %	7784	91,66 %	2613	96,43 %
Image1	7	82,23 %	12172	91,57 %	649	90,09 %
Mean		80,67 %		86,25 %		83,62 %

Représentation des connaissances



Ontologie des lettrines



Éléments de comparaison

Classification

- « Monde fermé »
- Apprentissage automatique du modèle
- Méthodes abouties (statistiques, RdF, bases de données, IA, ...)
- Données tabulaires « objets x attributs » numériques et symboliques
- Intègre la variabilité de données proches (mesure de distance/similarité)
- Possibilités d'interactivité « bas-niveau » (extraction de caractéristiques)

Représentation ontologique des connaissances

- « Monde ouvert »
- Construction du modèle (T-Box, A-Box)
- Emergence récente (web sémantique)
- Données complexes, attributaires, et relationnelles
- Mécanismes de raisonnement logique
- Possibilités d'interactivité niveau logique (règles, requêtes)

Conclusion

Combiner les approches de **classification** et de **représentation des connaissances** pour des **images de document**

- Images fortement structurées (information spatiale)
- Issues d'un domaine particulier (sémantique du domaine)

Images de
bandes
dessinées

Intérêt des **méthodes de l'AFC** (Analyse Formelle des Concepts)

- Classification & Représentation des connaissances
- Utilisation du jeu algorithmique adapté et efficace



Contributions et perspectives

2011 et après ?

Contributions

Lien algorithmique

Thème 1: Contributions structurelles et algorithmiques

- Aspects structurels
 - Graphe de dépendance*
 - Travaux de thèse*
- Usages d'un système de fermeture
 - Base Canonique Directe*
- Algorithmes de génération
 - Bibliothèque lattice*
 - Algorithmes autour d'un système de fermeture*

Contributions: algorithmes, structures, preuves

Thème 2: Quelques usages pour des données images

- Fouille de données
 - Méthode Navigala*
 - Treillis dichotomiques*
 - Logiciel Navigala*
- Relations spatiales et représentation des connaissances
 - Signature structurelle d'un symbole*
 - Ontologie des lettrines*

Contributions: méthodes, expérimentations

Diffusion scientifique

Thème 1: Contributions structurelles et algorithmiques

- Publications:
 - 6 revues
 - 9 conférences
- Bibliothèque lattice
- Projet:
 - PEPS 2010 (*Marseille, Paris, Caen, Angers, Lyon*)
- Animation:
 - Communauté francophone des treillis (*Paris, Nancy, Clermont-Ferrand, Montpellier, Marseille, Rennes, Montréal*)
 - Membre comité programme (CLA)
 - Relecture revues (INS, DKE)

Thème 2: Quelques usages pour des données images

- Publications:
 - 5 revues (2 nat.)
 - 9 conférences
 - 5 lncs
- Logiciel Navigala
- Encadrement:
 - 2 thèses
 - 8 stages M2
- Projets ANR:
 - Madonne (*Paris, Nancy, Rennes, Tours, Rouen*)
 - Navidomass (*Nancy, Tours Rouen*)
- Animation:
 - Equipe Imédoc puis IDDC
 - Consortium autour du document (*Nancy, Rouen, Tours, Barcelone, Paris*)
 - Membre comité programme (*GREC, CIFED*)

Perspectives

Lien algorithmique

Thème 1: Contributions structurelles et algorithmiques

- Aspects structurels
 - Liens entre graphe de dépendance et OD-Graph [Nat90,San09] (*projet PEPS 10*)
- Usages d'un système de fermeture
 - Bases de règles (BC, BCD, ...) (*projet PEPS 10*)
- Algorithmes de génération
 - Mise à disposition de la bibliothèque

Thème 2: Quelques usages pour des données images

- Fouille de données
 - Structure hybride de classification entre arbre et treillis (*thèse Nathalie Girard*)
 - Mécanisme d'explication par des règles (interactivité)
 - Indexation à l'aide de générateurs minimaux
- Relations spatiales et représentation des connaissances
 - Combiner information bas-niveau et information haut-niveau
 - Mécanismes d'interactivité
 - Cas d'étude: images de bandes dessinées (*projet CPER*) (*thèse Clément Guerin*)

Bibliographie

Thème 1: Contributions structurelles et algorithmiques

- **Treillis:** [GW99] [BM70] [MCo3] [Bir40] [Ded1903]
- **Extensions Galois:** [GK01] [GSo1] [FR04] [Pol98]
- **Bases:** [GD82] [RW95] [Wild94] [Mai83] [TBo2] [Mong0]
- **Algorithmes:** [Wild95] [Flo03] [Iba99] [Wild89] [Sho86] [Gor79] [Wil82] [Wil84] [Bor86]
- **Contributions structurelles:** [Bertet et al.02] [Bertet 98] [Bertet et al.99] [Bertet et al.97] [Bertet et al.00] [Bertet et al:Ordero2] [Bertet et al:Algebra01]
- **BCD:** [Bertet :SODA03] [Bertet et al:DMTCS04] [Bertet et al:TCS10]
- **Algorithmes:** [Bertet et al. 04] [Bertet et al.06] [Bertet11]

Thème 2: Quelques usages pour des données images

- **Classification:** [LV01] [LM90] [CR93] [Oos88] [Sah95] [NM99] [Xie 02] [UC10]
- **Ontologies et relations spatiales:** [Baa95] [Horo3] UK96] [BS01] [Sato8] [CBG97]
- **Analyse d'images:** [Hou59] [CO09] [DLo8] [GRECo3] [GRECo5] [Rad17] [TW03]
- **Méthode Navigala:** [Bertet et al.05] [Bertet et al.06] [Bertet et al.09] [Bertet et al.07] [Bertet et al:TS09] [Bertet et al.08] [Bertet et al.:IJPRAI10]
- **Signature structurelle:** [Bertet et al.:TS10] [Bertet et al.07] [Bertet et al.:SMC11]
- **Ontologie des lettrines:** [Bertet et al.:ICDAR11]



Merci de votre attention

Des questions ?