Des treillis, *d*-treillis, et quelques applications en analyse de données

Giacomo Kahn

12 décembre 2018

Sous la supervision d'Olivier Raynaud

Fromages d'Auvergne

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Table – Fromages d'Auvergne

Fromages d'Auvergne

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Table – Fromages d'Auvergne

Fromages d'Auvergne

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Table – Fromages d'Auvergne

Fromages d'Auvergne

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Table – Fromages d'Auvergne

On va appeler ces rectangles des **concepts**.

- Formalisation mathématique des notions sémantiques d'**intension** et d'**extension**.
- Interpretation du pragmatisme de Charles S. Pierce (1839-1914).

Fromages d'Auvergne

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Table – Fromages d'Auvergne

On va appeler ces rectangles des **concepts**.

- Formalisation mathématique des notions sémantiques d'**intension** et d'**extension**.
- Interpretation du pragmatisme de Charles S. Pierce (1839-1914).

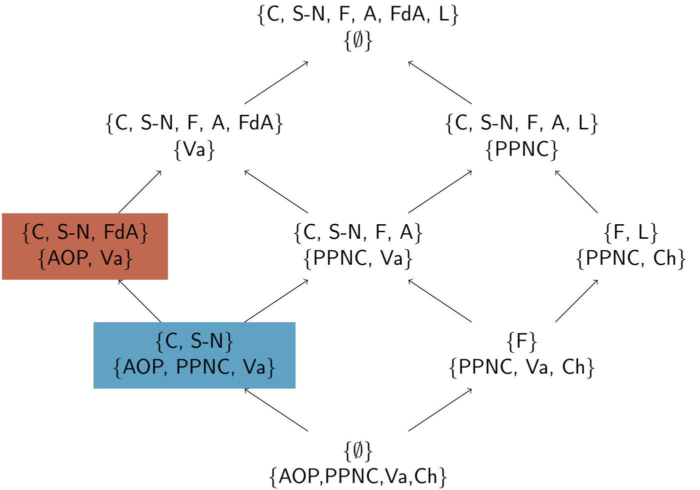


Figure – Treillis des concepts ($\mathcal{T}$(Fromages), $\subseteq_{\mathcal{O}}$)

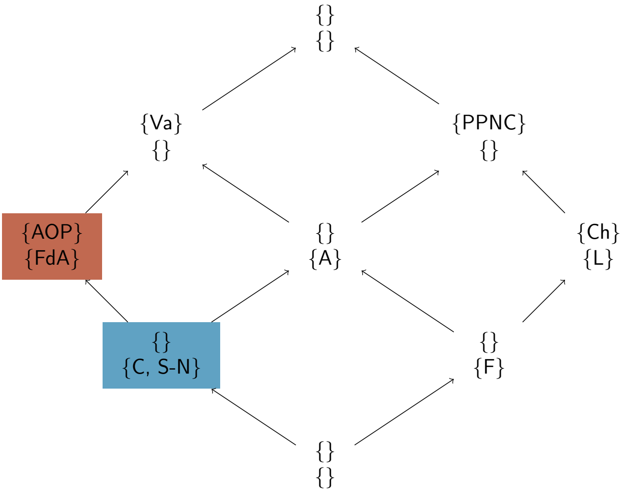


Figure – Treillis des concepts ($\mathcal{T}$(Fromages), $\subseteq_{\mathcal{O}}$)

L'Analyse Formelle de Concepts (FCA)

- Formalisée dans les années 80 à Darmstadt dans l'équipe de Rudolph Wille ;
- Utilise les outils mathématique de la théorie des treillis (Birkhoff, Barbut, Monjardet, et al.) pour l'analyse de données ;
- Nombreuses applications : représentation des connaissances, fouille de données, génie logiciel, etc.

Comment ça marche ?

FCA s'appuye sur une structure de initiale : le contexte.

Contexte

Un contexte est un triple $(\mathcal{O}, \mathcal{A}, \mathcal{R})$. $\mathcal{O}$ et $\mathcal{A}$ sont des ensembles, $\mathcal{R}$ est une relation entre eux.
Notre exemple présentait le contexte (Fromages, Caractéristiques, $\mathcal{R}$).

Comment ça marche ?

FCA s'appuye sur une structure de initiale : le contexte.

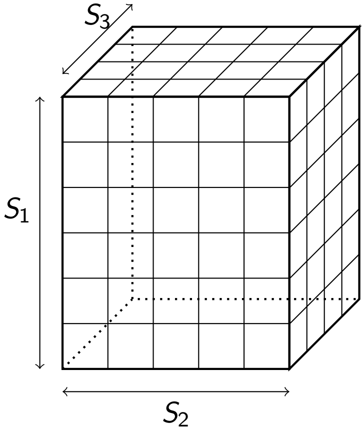
Contexte

Un contexte est un triple $(\mathcal{O}, \mathcal{A}, \mathcal{R})$. $\mathcal{O}$ et $\mathcal{A}$ sont des ensembles, $\mathcal{R}$ est une relation entre eux.
Notre exemple présentait le contexte (Fromages, Caractéristiques, $\mathcal{R}$).

- $(\cdot)'$: $\mathcal{O} \mapsto \mathcal{A}$: o un objet, o' représente l'ensemble d'attributs de cet objet.
- $(\cdot)'$: $\mathcal{A} \mapsto \mathcal{O}$: a un attribut, a' représente l'ensemble des objets qui ont cet attribut.

La composition $(\cdot)''$ est un opérateur de fermeture.

Passage en ~~trois~~ d dimensions

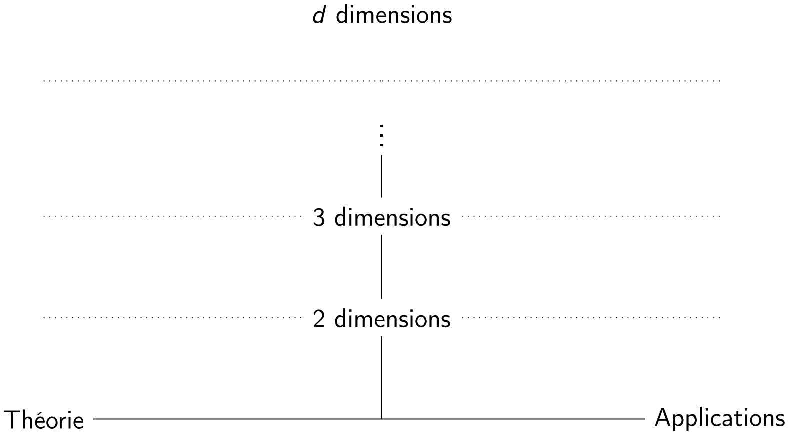


- Contexte $C = (\mathcal{S}_1, \ldots, \mathcal{S}_d, \mathcal{R})$,
- Concept : $(X_1, \ldots, X_d)$ boîte maximale d'incidence dans un contexte (boîte de croix/pavé combinatoire),
- Structure sous-jacente : d-treillis $(\mathcal{T}(C), \subseteq_{\mathcal{S}_1}, \ldots, \subseteq_{\mathcal{S}_d})$

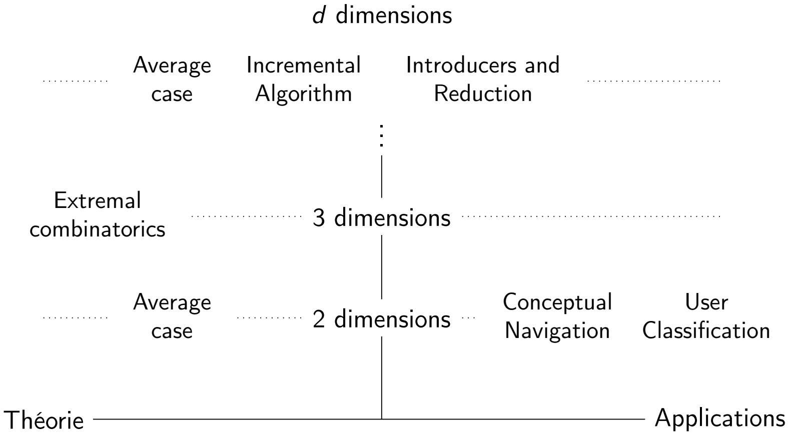
D'où viens-je ? Où vais-je ?

Théorie ——————————————— Applications

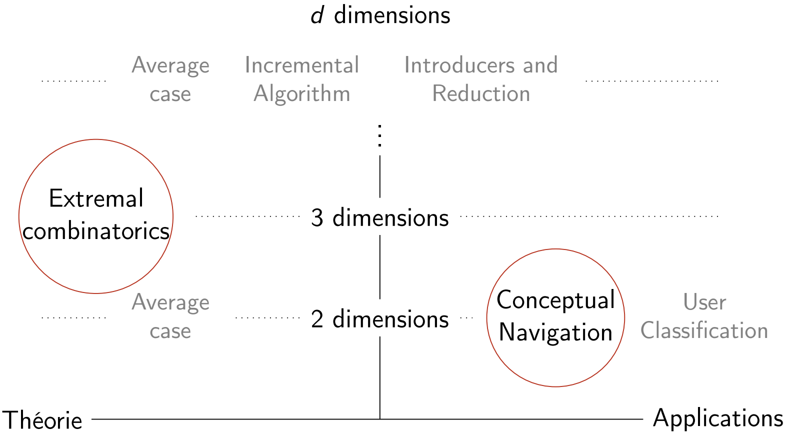
D'où viens-je ? Où vais-je ?



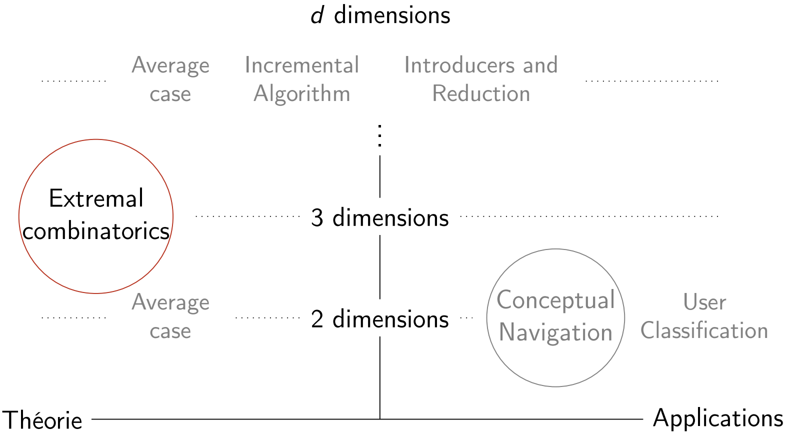
D'où viens-je ? Où vais-je ?



D'où viens-je ? Où vais-je ?



D'où viens-je ? Où vais-je ?



Le cas extrémal : nombre max de 2-concepts

- Un concept (X, Y) est défini de manière unique par une de ses composantes.

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Le cas extrémal : nombre max de 2-concepts

- Un concept (X, Y) est défini de manière unique par une de ses composantes.

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Le cas extrémal : nombre max de 2-concepts

- Un concept (X, Y) est défini de manière unique par une de ses composantes.
- Avec $|S_1| = |S_2| = s$, 2^s est une borne triviale.

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

Le cas extrémal : nombre max de 2-concepts

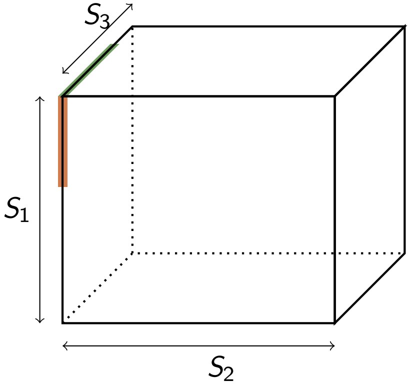
- Un concept (X, Y) est défini de manière unique par une de ses composantes.
- Avec $|S_1| = |S_2| = s$, 2^s est une borne triviale.

	1	2	3	4	5	6	7	8
1		×	×	×	×	×	×	×
2	×		×	×	×	×	×	×
3	×	×		×	×	×	×	×
4	×	×	×		×	×	×	×
5	×	×	×	×		×	×	×
6	×	×	×	×	×		×	×
7	×	×	×	×	×	×		×
8	×	×	×	×	×	×	×	

- Contranominal scale – échelle anti-diagonale $(S, S, \neq)$
- $2^{|S|}$ concepts : toutes les parties de S sont fermées (treillis booléen)

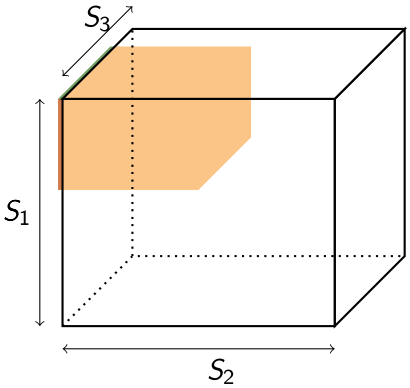
Le cas extrémal : nombre max de 3-concepts

- Un concept (X, Y, Z) est défini de manière unique par deux de ses trois composantes.



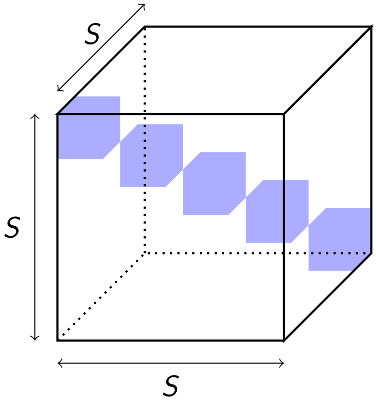
Le cas extrémal : nombre max de 3-concepts

- Un concept (X, Y, Z) est défini de manière unique par deux de ses trois composantes.
- Avec $|S_1| = |S_2| = |S_3| = s$, $2^s \times 2^s = 4^s$ est une borne triviale.



Le cas extrémal : nombre max de 3-concepts

- Un concept (X, Y, Z) est défini de manière unique par deux de ses trois composantes.
- Avec $|S_1| = |S_2| = |S_3| = s$, $2^s \times 2^s = 4^s$ est une borne triviale.



- Échelle anti-diagonale $(S, S, S, S^3 \setminus \{(a, a, a) \mid a \in S\})$
- $3^{|S|}$ concepts (3-treillis puissance)

Le cas extrémal : nombre max de 3-concepts

- Un concept (X, Y, Z) est défini de manière unique par deux de ses trois composantes.
- Avec $|S_1| = |S_2| = |S_3| = s$, $2^s \times 2^s = 4^s$ est une borne triviale.

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1		×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
2	×	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
3	×	×	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	×	×
4	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×
5	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
			1					2					3					4					5		

- Échelle anti-diagonale $(S, S, S, S^3 \setminus \{(a, a, a) \mid a \in S\})$
- $3^{|S|}$ concepts (3-treillis puissance)

Nombre maximum de 3-concepts

Question

Soit C un 3-contexte de taille $s \times s \times s$. Quel est le plus grand nombre de 3-concepts qui peuvent cohabiter dans C ?

Nombre maximum de 3-concepts

Question

Soit C un 3-contexte de taille $s \times s \times s$. Quel est le plus grand nombre de 3-concepts qui peuvent cohabiter dans C ?

$3^s \leq$ Nombre de 3-concepts $\leq 4^s$

Nombre maximum de 3-concepts

Question

Soit C un 3-contexte de taille $s \times s \times s$. Quel est le plus grand nombre de 3-concepts qui peuvent cohabiter dans C ?

$$3^s \leq \quad 3.36^s \leq \quad \text{Nombre de 3-concepts} \quad \leq 3.39^s \quad \leq 4^s$$

Construisons des contextes avec beaucoup de concepts

Jouons aux échecs

Comment placer 8 tours sur un échiquier de telle manière qu'elles ne puissent pas se prendre ?

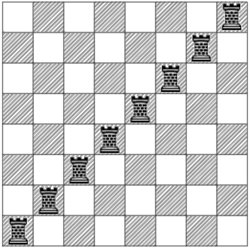


Figure – Solution : une tour par ligne et par colonne (image honteusement piquée sur http://mathworld.wolfram.com/RooksProblem.html)

Construisons des contextes avec beaucoup de concepts

Jouons aux échecs 3d

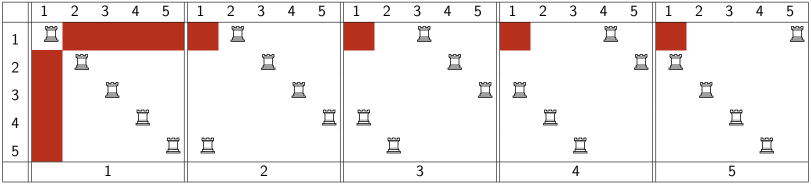
Plateau d'échec de $s \times s \times s$ avec des tours.

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	♖						♖						♖						♖						♖
2		♖						♖						♖						♖	♖				
3			♖						♖						♖	♖						♖			
4				♖						♖	♖						♖						♖		
5					♖	♖						♖						♖						♖	
			1					2					3					4					5		

Construisons des contextes avec beaucoup de concepts

Jouons aux échecs 3d

Plateau d'échec de $s \times s \times s$ avec des tours.



Construisons des contextes avec beaucoup de concepts

Jouons aux échecs 3d

Plateau d'échec de $s \times s \times s$ avec des tours.

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	♖						♖						♖						♖						♖
2		♖						♖						♖						♖	♖				
3			♖						♖						♖	♖						♖			
4				♖						♖	♖						♖						♖		
5					♖	♖						♖						♖						♖	
	1					2					3					4					5				

Construisons des contextes avec beaucoup de concepts

Jouons aux échecs 3d

Plateau d'échec de $s \times s \times s$ avec des trous.

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×	
2	×		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×			×	×	×	×
3	×	×		×	×	×	×	×		×	×	×	×	×			×	×	×	×	×		×	×	×
4	×	×	×		×	×	×	×	×			×	×	×	×	×		×	×	×	×	×		×	×
5	×	×	×	×			×	×	×	×	×		×	×	×	×	×		×	×	×	×	×		×
			1					2					3					4					5		

Construisons des contextes avec beaucoup de concepts

Jouons aux échecs 3d

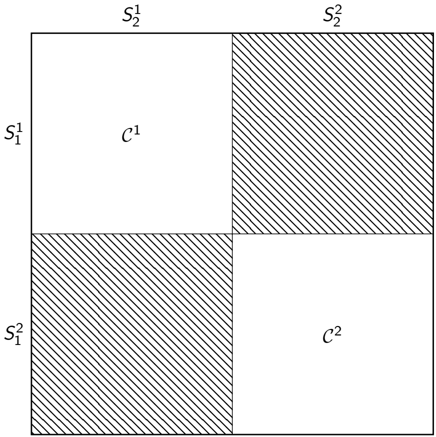
Plateau d'échec de $s \times s \times s$ avec des trous.

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×	
2	×		×	×	×	×	×		×	×	×	×	×		×	×	×	×	×			×	×	×	×
3	×	×		×	×	×	×	×		×	×	×	×	×			×	×	×	×	×		×	×	×
4	×	×	×		×	×	×	×	×			×	×	×	×	×		×	×	×	×	×		×	×
5	×	×	×	×			×	×	×	×	×		×	×	×	×	×		×	×	×	×	×		×
			1					2					3					4					5		

Gagné ?

- Solutions au problème des tours : 3.10^3, 3.16^4.
- Ce contexte a 3.36^5 concepts.
- Mais ! Au dessus de $s = 5$, les solutions du problèmes des tours en 3d donnent moins de 3^s concepts.

Une multiplication salvatrice

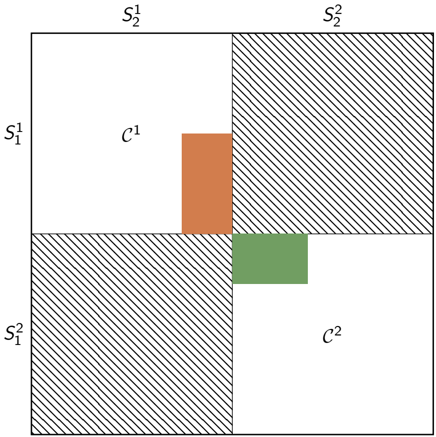


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Une multiplication salvatrice

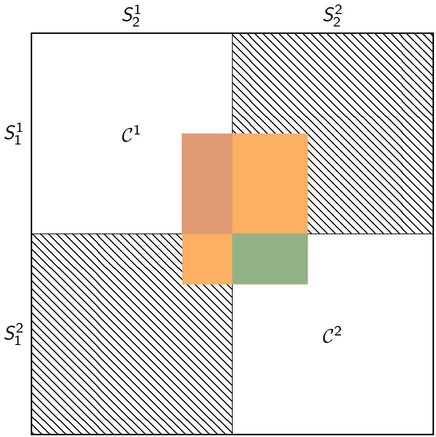


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Une multiplication salvatrice

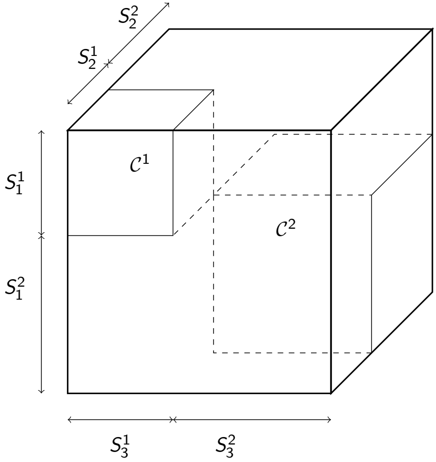


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Une multiplication salvatrice

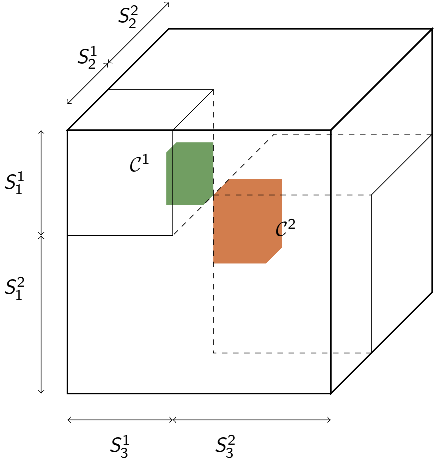


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Une multiplication salvatrice

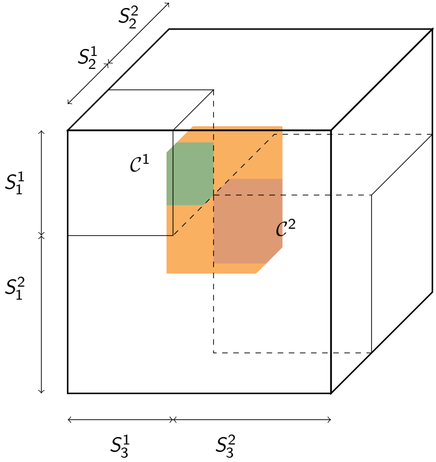


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Une multiplication salvatrice

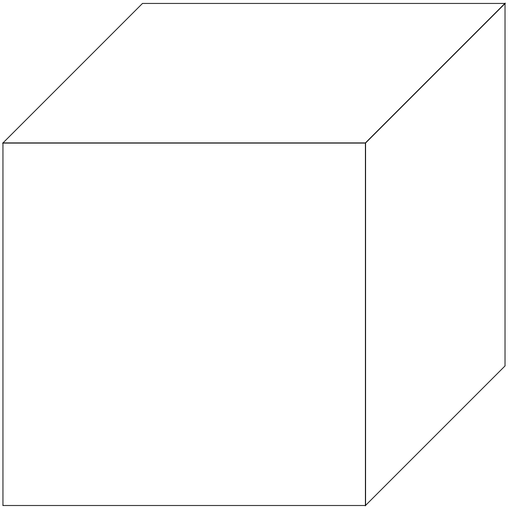


- Union des dimensions,
- union des relations,
- ajout des croix dans les cellules hybrides.

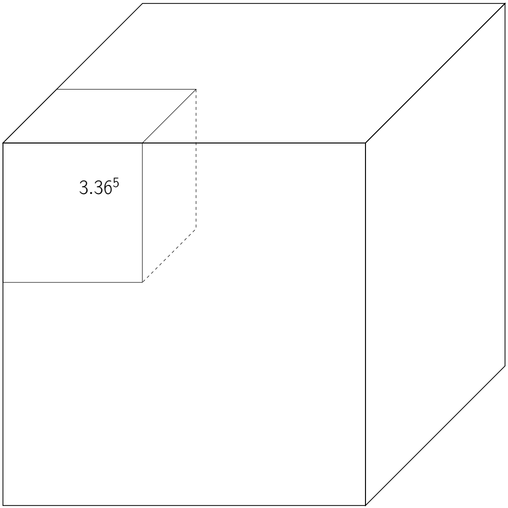
Cela entraîne la multiplication des concepts

$$|\mathcal{T}(C)| = |\mathcal{T}(\mathcal{C}^1)| \times |\mathcal{T}(\mathcal{C}^2)|$$

Exemple : $13 \times 13 \times 13$

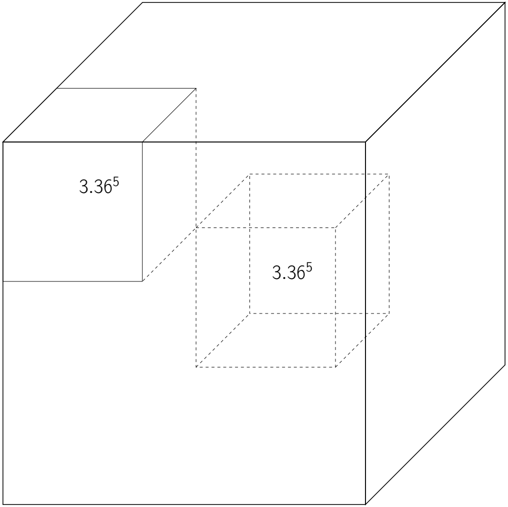


Exemple : $13 \times 13 \times 13$



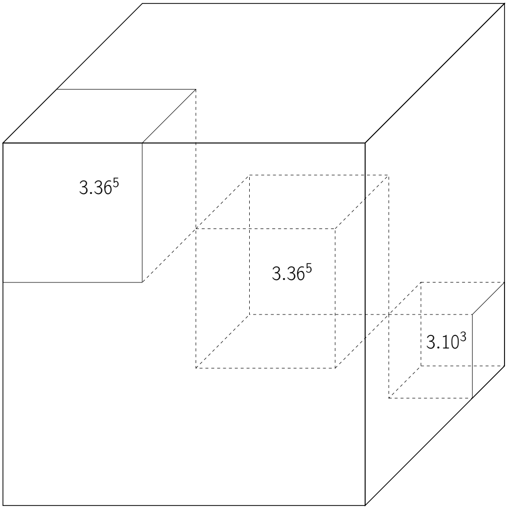
3.36^5 concepts.

Exemple : $13 \times 13 \times 13$



3.36^{10} concepts.

Exemple : $13 \times 13 \times 13$



$(\frac{3.10}{3.36})^3 \times 3.36^{13}$ concepts.

On peut maintenant atteindre 3.36^s

Multiplication de contexte

Solution du problème des tours + multiplication de contextes = on atteint $c3.36^s$.

La constante c provient du résultat de s modulo 5 (on doit parfois compléter avec de plus petits contextes).

On peut maintenant atteindre 3.36^s

Multiplication de contexte

Solution du problème des tours + multiplication de contextes = on atteint $c3.36^s$.
La constante c provient du résultat de s modulo 5 (on doit parfois compléter avec de plus petits contextes).

Question

Soit C un 3-contexte de taille $s \times s \times s$. Quel est le plus grand nombre de 3-concepts qui peuvent cohabiter dans C ?

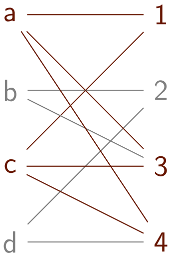
Réponse

- Plus que 3.36^s, ✓
- Moins que 3.39^s.

Transformation : des concepts aux transversaux

- Contexte = graphe biparti,
- Croix $\rightarrow$ arête,
- Concept $\rightarrow$ bi-clique max.

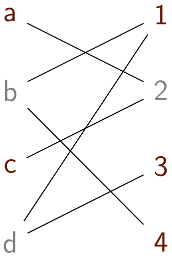
	a	b	c	d
1	×		×	
2		×		×
3	×	×	×	
4	×		×	×



Transformation : des concepts aux transversaux

- Contexte = graphe biparti,
- Trou $\rightarrow$ arête,
- Concept $\rightarrow$ stable max.

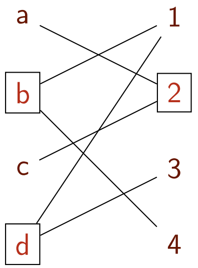
	a	b	c	d
1	$\times$		$\times$	
2		$\times$		$\times$
3	$\times$	$\times$	$\times$	
4	$\times$		$\times$	$\times$



Transformation : des concepts aux transversaux

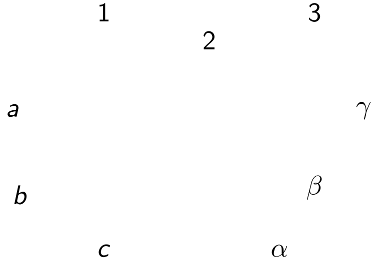
- Contexte = graphe biparti,
- Trou $\rightarrow$ arête,
- Concept $\rightarrow$ $\overline{\text{stable max}}$ $\rightarrow$ transversal minimal

	a	b	c	d
1	×		×	
2		×		×
3	×	×	×	
4	×		×	×



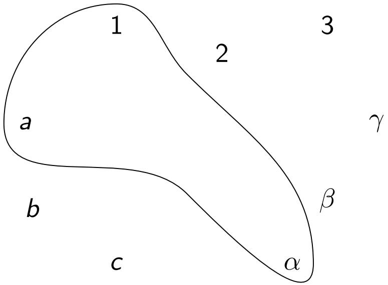
Transformation : des concepts aux transversaux

		1	2	3
α	a		×	×
	b	×	×	×
	c	×	×	×
β	a	×	×	×
	b	×		×
	c	×	×	×
γ	a	×	×	×
	b	×	×	×
	c	×	×	



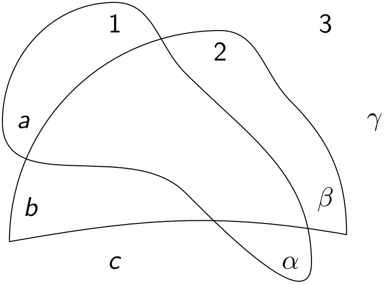
Transformation : des concepts aux transversaux

		1	2	3
α	a		×	×
	b	×	×	×
	c	×	×	×
β	a	×	×	×
	b	×		×
	c	×	×	×
γ	a	×	×	×
	b	×	×	×
	c	×	×	



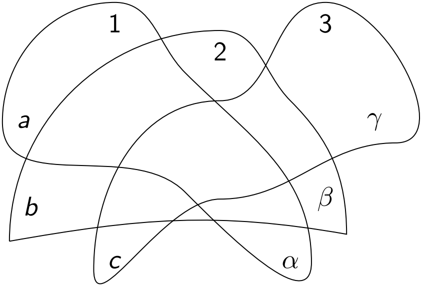
Transformation : des concepts aux transversaux

		1	2	3
α	a		×	×
	b	×	×	×
	c	×	×	×
β	a	×	×	×
	b	×		×
	c	×	×	×
γ	a	×	×	×
	b	×	×	×
	c	×	×	



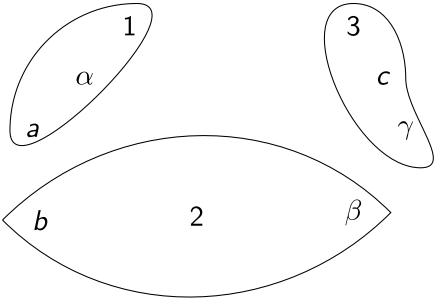
Transformation : des concepts aux transversaux

		1	2	3
α	a		×	×
	b	×	×	×
	c	×	×	×
β	a	×	×	×
	b	×		×
	c	×	×	×
γ	a	×	×	×
	b	×	×	×
	c	×	×	



Transformation : des concepts aux transversaux

		1	2	3
α	a		×	×
	b	×	×	×
	c	×	×	×
β	a	×	×	×
	b	×		×
	c	×	×	×
γ	a	×	×	×
	b	×	×	×
	c	×	×	



À quoi ressemblent nos hypergraphes ?

Contexte $s \times s \times s \rightarrow$ hypergraphe à $n = 3 \times s$ sommets.

Une classe contrainte

- 3 dimensions $\rightarrow$ 3 ensembles indépendants ;
- 3 coordonnées $\rightarrow$ toutes les arêtes sont de taille 3 ;
- chacune des arête emprunte une coordonnée à toutes les dimensions.

Classe des hypergraphes 3-partite 3-uniformes.

Équivalence : les bornes

$$3^s \leq \text{Nombre max de 3-concepts} \leq 4^s.$$

Avec $n = 3s$:

$$1.4422^n \leq \text{Nombre max de transversaux min} \leq 1.5875^n.$$

Équivalence : les bornes

$$3.36^s \leq \text{Nombre max de 3-concepts} \leq 4^s.$$

Avec $n = 3s$:

$$1.4977^n \leq \text{Nombre max de transversaux min} \leq 1.5875^n.$$

Équivalence : les bornes

$$3.36^s \leq \text{Nombre max de 3-concepts} \leq 4^s.$$

Avec $n = 3s$:

$$1.4977^n \leq \text{Nombre max de transversaux min} \leq 1.5875^n.$$

Theorème (Lonc et Truszczyński 2008)

Le nombre maximum de transversaux minimaux dans les hypergraphes 3-uniformes à n sommets est borné par 1.67^n.

Équivalence : les bornes

$$3.36^s \leq \text{Nombre max de 3-concepts} \leq 4^s.$$

Avec $n = 3s$:

$$1.4977^n \leq \text{Nombre max de transversaux min} \leq 1.5875^n.$$

Theorème (Lonc et Truszczyński 2008)

Le nombre maximum de transversaux minimaux dans les hypergraphes 3-uniformes à n sommets est borné par 1.67^n.

Peut-on utiliser la même approche qu'eux ?

Oui. ☺

Équivalence : les bornes

$$3.36^s \leq \text{Nombre max de 3-concepts} \leq 4^s.$$

Avec $n = 3s$:

$$1.4977^n \leq \text{Nombre max de transversaux min} \leq 1.5875^n.$$

Theorème (Lonc et Truszczyński 2008)

Le nombre maximum de transversaux minimaux dans les hypergraphes 3-uniformes à n sommets est borné par 1.67^n.

Peut-on utiliser la même approche qu'eux ?

Oui. ☺ Approche : (Kullmann 1999)

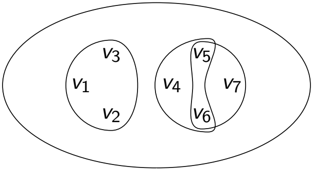
Algorithme

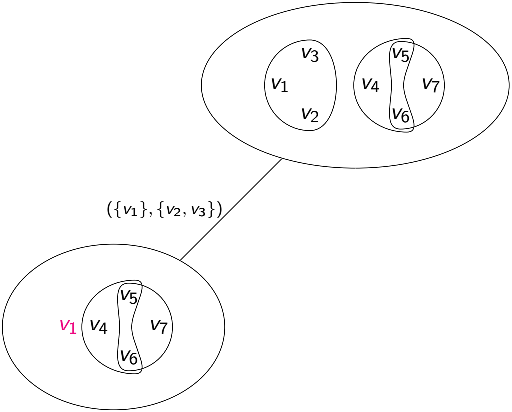
Calcul des transversaux min d'un hypergraphe

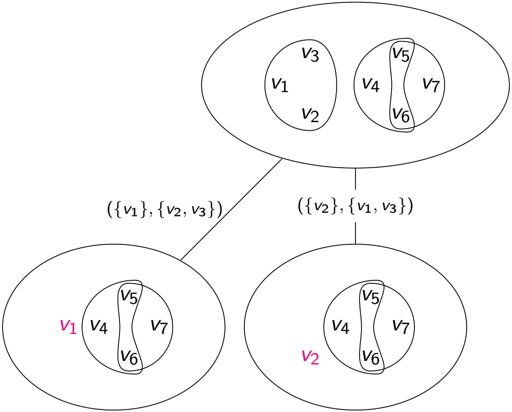
Approche classique : on fixe une partie de la solution et on continue sur une instance plus petite.

Prise de décisions

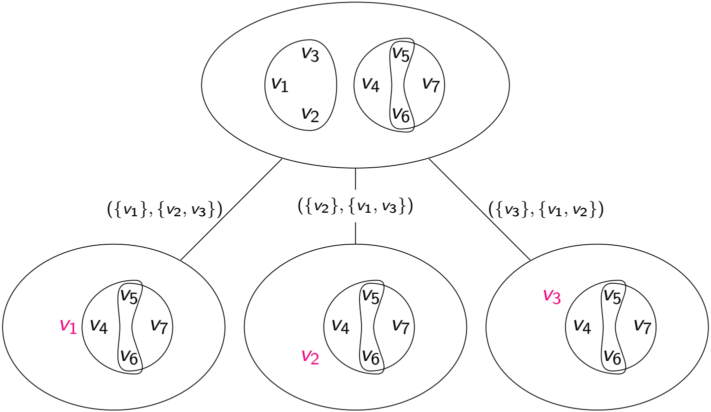
Déterminer un transversal min, c'est déterminer pour chaque sommet de l'hypergraphe si il fait partie du transversal.

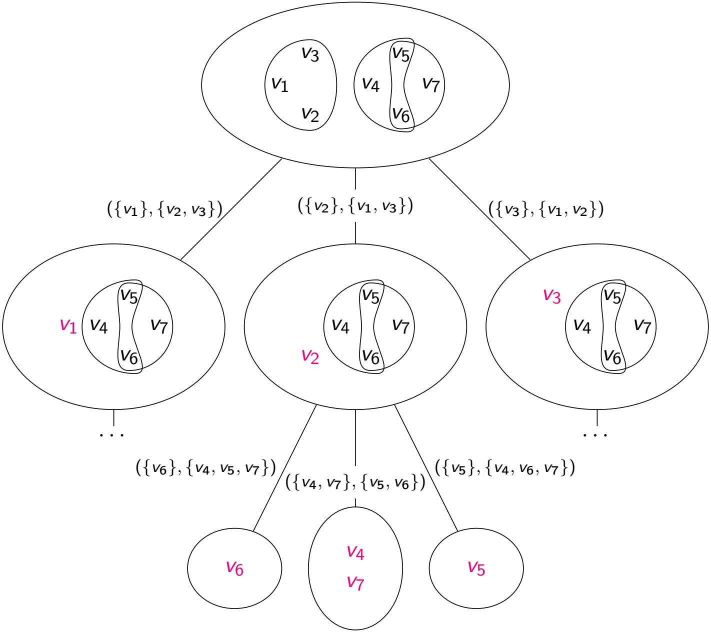

v_3
v_1
v_2
v_5
v_4
v_7
v_6


v_3
v_1
v_2
v_5
v_4
v_7
v_6
$(\{v_1\}, \{v_2, v_3\})$
v_5
v_1
v_4
v_7
v_6



v_3
v_1
v_2
v_5
v_4
v_7
v_6
$(\{v_1\}, \{v_2, v_3\})$
$(\{v_2\}, \{v_1, v_3\})$
v_5
v_1
v_4
v_7
v_6
v_5
v_4
v_7
v_2
v_6


v3
v1
v2
v5
v4
v7
v6
({v1}, {v2, v3})
({v2}, {v1, v3})
({v3}, {v1, v2})
v1 v4 v5 v7 v6
v4 v5 v7 v2 v6
v3 v4 v5 v7 v6

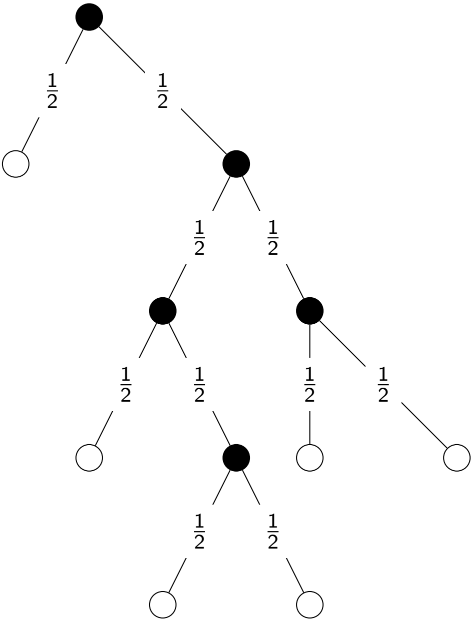

v_3
v_1
v_2
v_5
v_4
v_7
v_6
$(\{v_1\},\{v_2,v_3\})$
$(\{v_2\},\{v_1,v_3\})$
$(\{v_3\},\{v_1,v_2\})$
v_1
v_2
v_3
...
...
$(\{v_6\},\{v_4,v_5,v_7\})$
$(\{v_4,v_7\},\{v_5,v_6\})$
$(\{v_5\},\{v_4,v_6,v_7\})$
v_6
v_4
v_7
v_5

Nouvelle mission

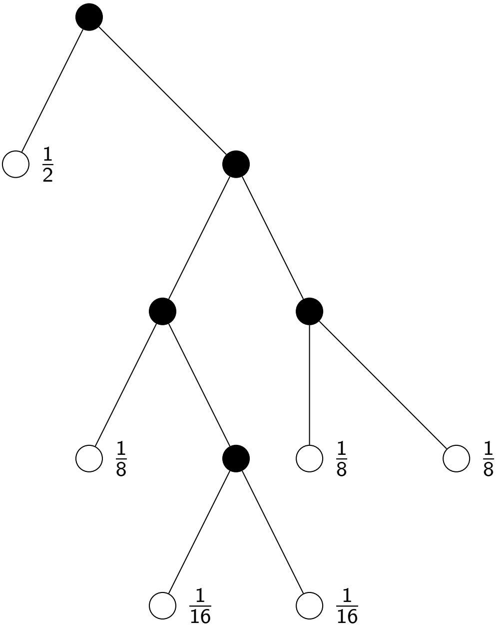
Théorème (Kullmann 1999)

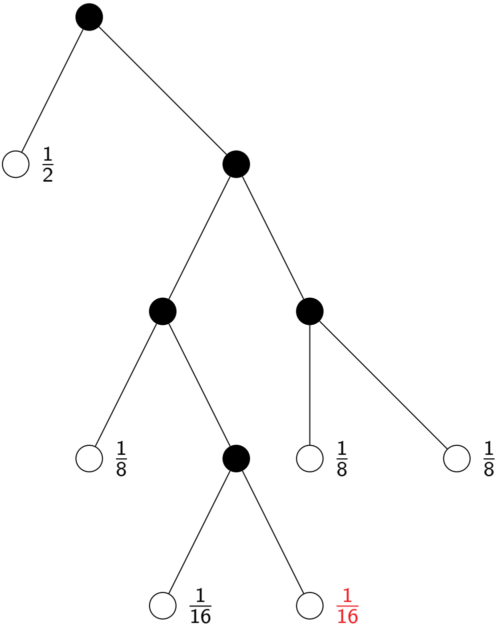
Un arbre enraciné muni d'une probabilité de transition a son nombre de feuilles borné par :

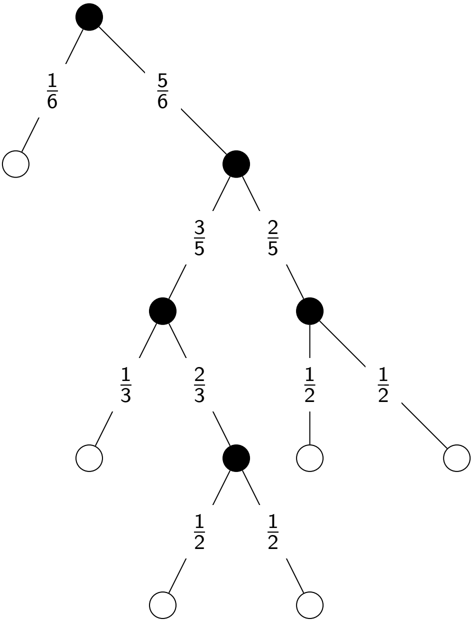
$$\text{Nb feuilles} \leq \max_{\ell \in L(\mathcal{T})} \left(\prod_{e \in P(\ell)} p(e) \right)^{-1}$$

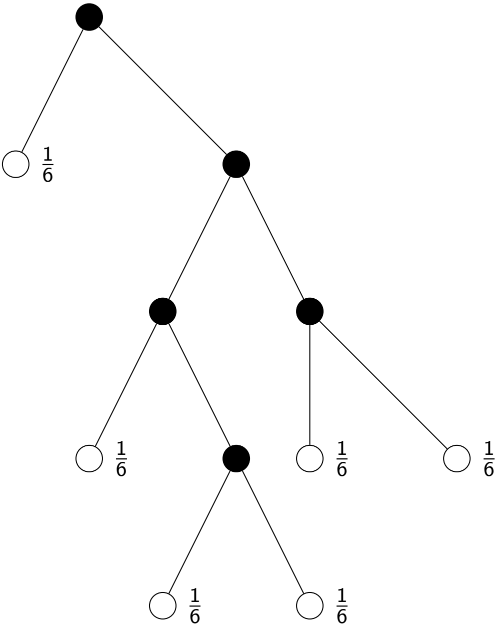


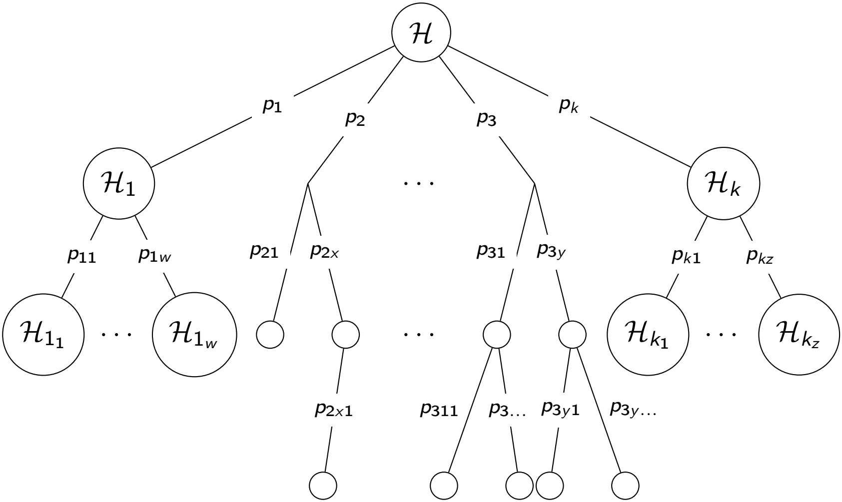
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$

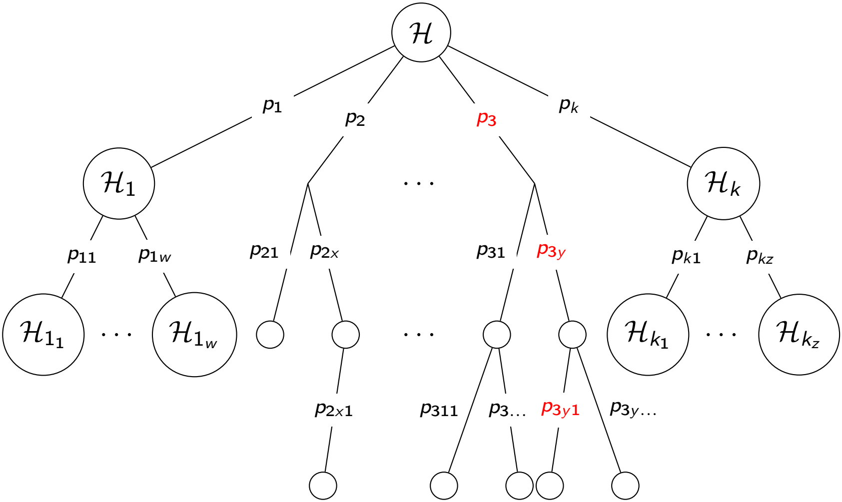

$\frac{1}{2}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{16}$
$\frac{1}{16}$


$\frac{1}{2}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{16}$
$\frac{1}{16}$


$\frac{1}{6}$
$\frac{5}{6}$
$\frac{3}{5}$
$\frac{2}{5}$
$\frac{1}{3}$
$\frac{2}{3}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$
$\frac{1}{2}$


$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$


$\mathcal{H}$
p_1
p_2
p_3
p_k
$\mathcal{H}_1$
$\dots$
$\mathcal{H}_k$
p_{11}
p_{1w}
p_{21}
p_{2x}
p_{31}
p_{3y}
p_{k1}
p_{kz}
$\mathcal{H}_{1_1}$
$\dots$
$\mathcal{H}_{1_w}$
$\dots$
$\mathcal{H}_{k_1}$
$\dots$
$\mathcal{H}_{k_z}$
p_{2x1}
p_{311}
$p_{3\dots}$
p_{3y1}
$p_{3y\dots}$


$\mathcal{H}$
p_1
p_2
p_3
p_k
$\mathcal{H}_1$
$\cdots$
$\mathcal{H}_k$
p_{11}
p_{1w}
p_{21}
p_{2x}
p_{31}
p_{3y}
p_{k1}
p_{kz}
$\mathcal{H}_{1_1}$
$\cdots$
$\mathcal{H}_{1_w}$
$\cdots$
$\mathcal{H}_{k_1}$
$\cdots$
$\mathcal{H}_{k_z}$
p_{2x1}
p_{311}
$p_{3\ldots}$
p_{3y1}
$p_{3y\ldots}$

Fin de la preuve

Ajuster la probabilité en fonction de paramètres structurels de l'hypergraphe.

Lemme (Bazin, Beaudou, K., Khoshkhah)

Il existe un ensemble de règles de branchement tel que le nombre de transversaux min d'un hypergraphe 3-partite 3-uniformes à n sommets est borné par $1.8393^{\frac{2n}{3}}$.

Fin de la preuve

Ajuster la probabilité en fonction de paramètres structurels de l'hypergraphe.

Lemme (Bazin, Beaudou, K., Khoshkhah)

Il existe un ensemble de règles de branchement tel que le nombre de transversaux min d'un hypergraphe 3-partite 3-uniformes à n sommets est borné par $1.8393^{\frac{2n}{3}}$.

$$1.8393^{\frac{2n}{3}} = 1.50119^{n} \leq 3.39^{s}$$

Fin de la preuve

Ajuster la probabilité en fonction de paramètres structurels de l'hypergraphe.

Lemme (Bazin, Beaudou, K., Khoshkhah)

Il existe un ensemble de règles de branchement tel que le nombre de transversaux min d'un hypergraphe 3-partite 3-uniformes à n sommets est borné par $1.8393^{\frac{2n}{3}}$.

$$1.8393^{\frac{2n}{3}} = 1.50119^n \leq 3.39^s$$

$$3^s \leq \quad 3.36^s \leq \quad \text{Nombre de 3-concepts} \quad \leq 3.39^s \quad \leq 4^s$$

Le nombre de concepts est très exponentiel

Théorème

Le nombre maximum de 3-concepts dans un 3-contexte de taille $s \times s \times s$ est au moins 3.36^s.

Théorème

Le nombre maximum de 3-concepts dans un 3-contexte de taille $s \times s \times s$ est au plus 3.39^s.

Le nombre de concepts est très exponentiel

Théorème

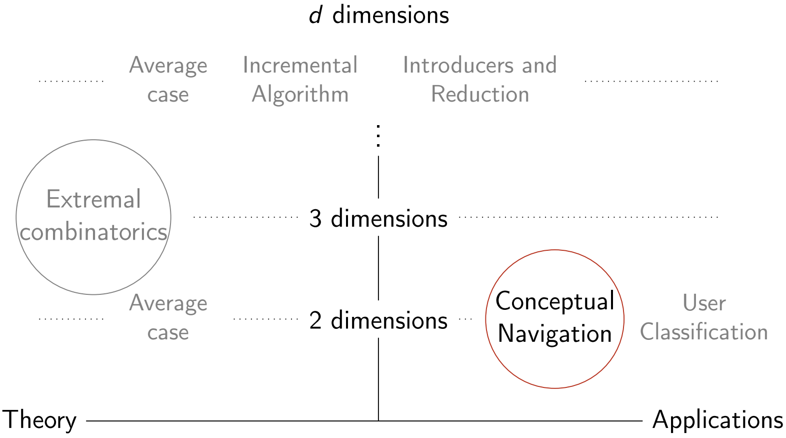
Le nombre maximum de 3-concepts dans un 3-contexte de taille $s \times s \times s$ est au moins 3.36^s.

Théorème

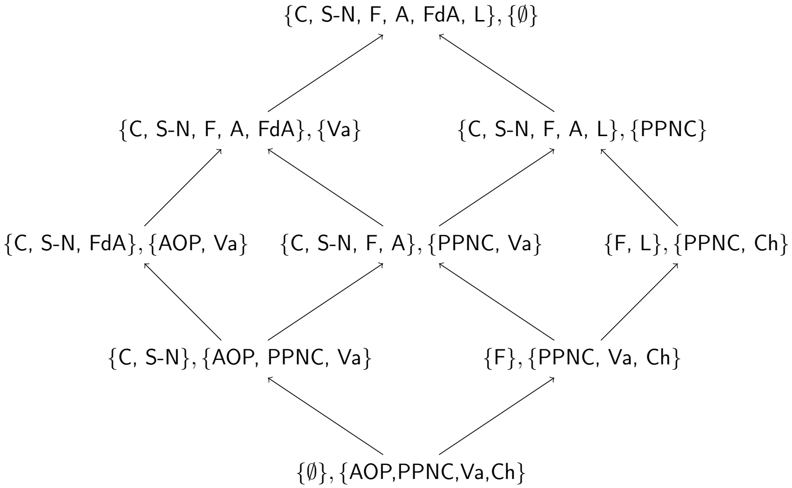
Le nombre maximum de 3-concepts dans un 3-contexte de taille $s \times s \times s$ est au plus 3.39^s.

Que faire ?

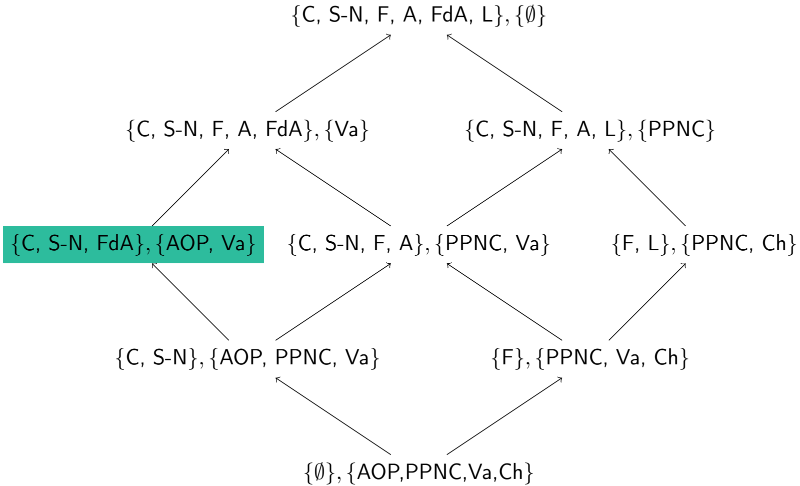
- Algorithmes incrémentaux ;
- Réduction des structures (bases d'implications, sous-structures gardant l'information) ;
- Approches locales.


d dimensions
Average case
Incremental Algorithm
Introducers and Reduction
Extremal combinatorics
3 dimensions
Average case
2 dimensions
Conceptual Navigation
User Classification
Theory
Applications

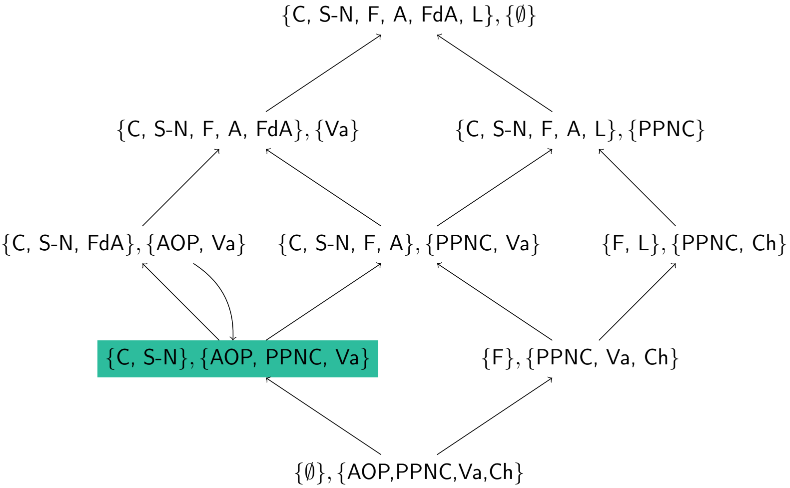
Navigation conceptuelle



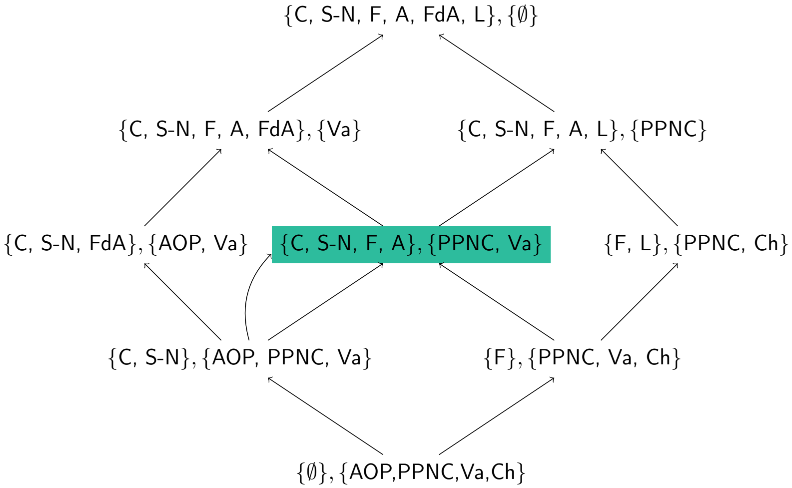
Navigation conceptuelle



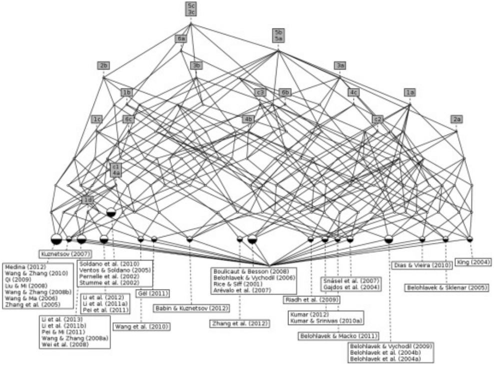
Navigation conceptuelle



Navigation conceptuelle

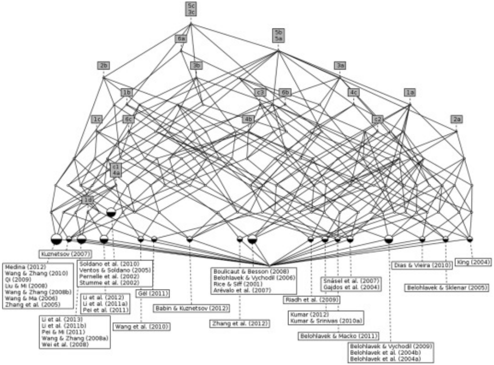


Les treillis de concepts sont vite grands



- Stockage ;
- orientation ;
- concepts "inutiles".

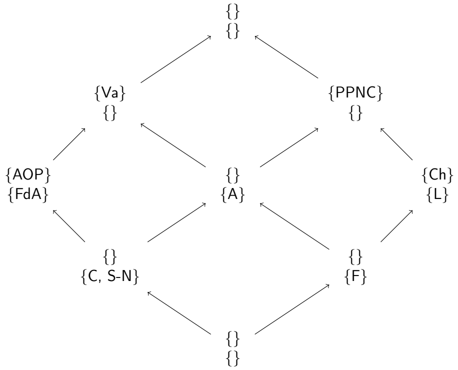
Les treillis de concepts sont vite grands



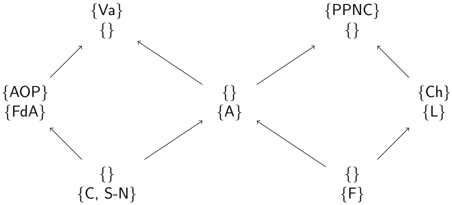
- Stockage ;
- orientation ;
- concepts "inutiles".

On voudrait une structure plus petite.

Concepts introducteurs



Concepts introducteurs



Avantage : taille linéaire

On a au plus un concept pour chacun des objets et pour chacun des attributs.

“Certaines applications”

Oui mais lesquelles

- Objet : identifiant d’un ensemble d’attribut.
- Pas de valeur sémantique, juste symbolique.

“Certaines applications”

Oui mais lesquelles

- Objet : identifiant d’un ensemble d’attribut.
- Pas de valeur sémantique, juste symbolique.

Exemples

- Génie logiciel : produit logiciel = ensemble de descripteurs.
- Classification d’utilisateurs : session = ensemble d’actions.

Navigation conceptuelle dans les AOC-posets

Contribution

Algorithme de navigation conceptuelle dans un AOC-poset, à partir de n'importe quel configuration d'attributs (ou d'objets).

Navigation conceptuelle dans les AOC-posets

Contribution

Algorithme de navigation conceptuelle dans un AOC-poset, à partir de n'importe quel configuration d'attributs (ou d'objets).

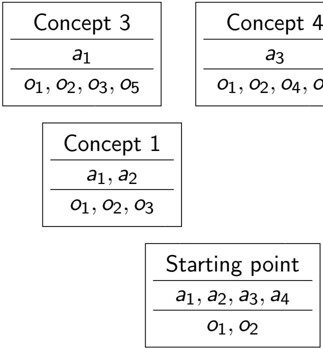
Starting point
a_1, a_2, a_3, a_4
o_1, o_2

	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×	×	×
o_3	×	×		
o_4			×	×
o_5	×			
o_6			×	
o_7				×

Navigation conceptuelle dans les AOC-posets

Contribution

Algorithme de navigation conceptuelle dans un AOC-poset, à partir de n'importe quel configuration d'attributs (ou d'objets).



	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×	×	×
o_3	×	×		
o_4			×	×
o_5	×			
o_6			×	
o_7				×

Navigation conceptuelle dans les AOC-posets

Contribution

Algorithme de navigation conceptuelle dans un AOC-poset, à partir de n'importe quel configuration d'attributs (ou d'objets).

Concept 3
a_1
o_1, o_2, o_3, o_5

Concept 4
a_3
o_1, o_2, o_4, o_6

Concept 5
a_4
o_1, o_2, o_4, o_7

Concept 1
a_1, a_2
o_1, o_2, o_3

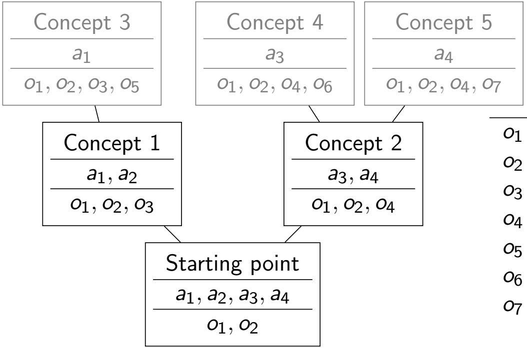
Starting point
a_1, a_2, a_3, a_4
o_1, o_2

	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×	×	×
o_3	×	×		
o_4			×	×
o_5	×			
o_6			×	
o_7				×

Navigation conceptuelle dans les AOC-posets

Contribution

Algorithme de navigation conceptuelle dans un AOC-poset, à partir de n'importe quel configuration d'attributs (ou d'objets).



	a_1	a_2	a_3	a_4
o_1	×	×	×	×
o_2	×	×	×	×
o_3	×	×		
o_4			×	×
o_5	×			
o_6			×	
o_7				×

Génération locale

Le point de départ n'est pas défini ; on peut donc générer toute la structure de cette manière.

Performances

Plus le treillis des concepts est grands, plus il est avantageux d'utiliser l'AOC-poset. La génération locale gagne aussi en performances à ce niveau là.

Génération locale

Le point de départ n'est pas défini ; on peut donc générer toute la structure de cette manière.

Performances

Plus le treillis des concepts est grands, plus il est avantageux d'utiliser l'AOC-poset. La génération locale gagne aussi en performances à ce niveau là.

Peut-on étendre cette approche ?

On souhaiterait étendre l'approche à la navigation conceptuelle dans une famille de contextes RCA.

Analyse relationnelle de concepts

Insérer des relations dans FCA

Une famille relationnelle de contextes est une paire $(\mathbf{K}, \mathbf{R})$ où $\mathbf{K}$ et un ensemble de contextes et $\mathbf{R}$ est un ensemble de relations entre les objets de contextes de $\mathbf{K}$.

Exemple

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

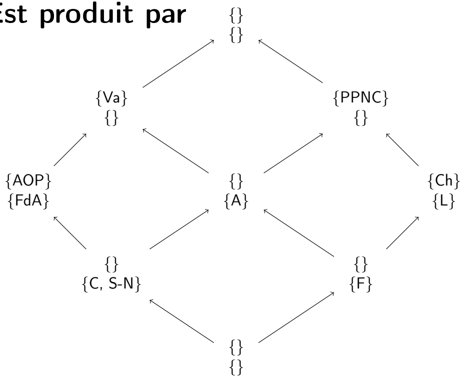
	Bio	Local	Vente directe
Dischamp		×	
Laqueuille	×		×
Roger	×	×	×
Emile	×	×	×

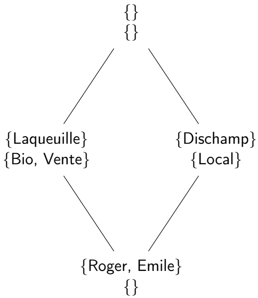
Est produit par	Dischamp	Laqueuille	Roger	Emile
Cantal	×			
Saint-Nectaire	×		×	
Fourme d'Ambert	×	×		
Fouchtra			×	
Lavort				×
Artisous				×

Concepts puis relations

Quantifieur ∃

On étend le contexte de départ si il existe un objet qui intersecte l'extention d'un des concepts du contexte cible. Exemple : relation **Est produit par**



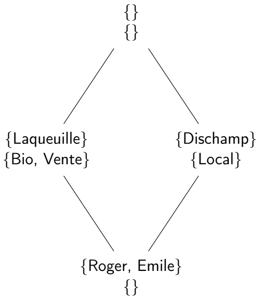


Concepts puis relations

Quantifieur ∃

On étend le contexte de départ si il existe un objet qui intersecte l'extention d'un des concepts du contexte cible. Exemple : relation **Est produit par**

	AOP	Pâte pressée non cuite	Vache	Chèvre
Cantal	×	×	×	
Saint-Nectaire	×	×	×	
Fourme d'Ambert	×		×	
Fouchtra		×	×	×
Lavort		×		×
Artisous		×	×	

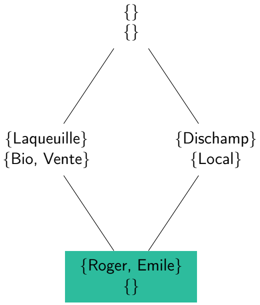


Concepts puis relations

Quantifieur ∃

On étend le contexte de départ si il existe un objet qui intersecte l'extention d'un des concepts du contexte cible. Exemple : relation **Est produit par**

	AOP	Pâte pressée non cuite	Vache	Chèvre	∃Est Produit Par {RE, Bio Local Vente}
Cantal	×	×	×		
Saint-Nectaire	×	×	×		×
Fourme d'Ambert	×		×		
Fouchtra		×	×	×	×
Lavort		×		×	×
Artisous		×	×		×



Concepts puis relations

Quantifieur ∃

On étend le contexte de départ si il existe un objet qui intersecte l'extention d'un des concepts du contexte cible. Exemple : relation **Est produit par**

	AOP	Pâte pressée non cuite	Vache	Chèvre	∃Est Produit Par {RE, Bio Local Vente}	∃Est Produit Par {LRE, Bio Vente}	∃Est Produit Par {DRE, Local }	∃Est Produit Par {DLRE, }
Cantal	×	×	×				×	×
Saint-Nectaire	×	×	×		×	×	×	×
Fourme d'Ambert	×		×			×	×	×
Fouchtra		×	×	×	×	×	×	×
Lavort		×		×	×	×	×	×
Artisous		×	×		×	×	×	×

Concepts puis relations

Quantifieur ∃

On étend le contexte de départ si il existe un objet qui intersecte l'extention d'un des concepts du contexte cible. Exemple : relation **Est produit par**

	AOP	Pâte pressée non cuite	Vache	Chèvre	∃**Est Produit Par** {**RE**, **Bio Local Vente**}	∃Est Produit Par {LRE, Bio Vente}	∃Est Produit Par {DRE, Local }	∃Est Produit Par {DLRE, }
Cantal	×	×	×				×	×
Saint-Nectaire	×	×	×		×	×	×	×
Fourme d'Ambert	×		×			×	×	×
Fouchtra		×	×	×	×	×	×	×
Lavort		×		×	×	×	×	×
Artisous		×	×		×	×	×	×

Et on recommence

On continue ce processus jusqu'à un **point fixe**.

- Calcul des treillis de concepts ;
- ajout des attributs relationnels.

Un point fixe est atteint lorsque les treillis obtenus à l'étape $\ell + 1$ sont isomorphes à ceux de l'étape ℓ.

Et on recommence

On continue ce processus jusqu'à un **point fixe**.

- Calcul des treillis de concepts ;
- ajout des attributs relationnels.

Un point fixe est atteint lorsque les treillis obtenus à l'étape $\ell + 1$ sont isomorphes à ceux de l'étape ℓ.

Inconvénient

On calcule de nombreuses fois des treillis de concepts.

Et on recommence

On continue ce processus jusqu'à un **point fixe**.

- Calcul des treillis de concepts ;
- ajout des attributs relationnels.

Un point fixe est atteint lorsque les treillis obtenus à l'étape $\ell + 1$ sont isomorphes à ceux de l'étape ℓ.

Inconvénient

On calcule de nombreuses fois des treillis de concepts.

Concept introducteurs à la rescousse

On peut remplacer les treillis par les AOC-posets correspondant.

Navigation conceptuelle dans RCA

Objectif

Point de départ : un concept.

On veut :

- sa couverture supérieure ;
- sa couverture inférieure ;
- sa couverture relationnelle.

Navigation conceptuelle dans RCA

Objectif

Point de départ : un concept.
On veut :

- sa couverture supérieure ;
- sa couverture inférieure ;
- sa couverture relationnelle.

Les concepts immédiatement au dessus dans l'AOC-poset

Navigation conceptuelle dans RCA

Objectif

Point de départ : un concept.
On veut :

- sa couverture supérieure ;
- sa couverture inférieure ;
- sa couverture relationnelle.

Les concepts immédiatement au dessus dans l'AOC-poset

Les concepts immédiatement en dessous dans l'AOC-poset

Navigation conceptuelle dans RCA

Objectif

Point de départ : un concept.
On veut :

- sa couverture supérieure ;
- sa couverture inférieure ;
- sa couverture relationnelle.

Les concepts immédiatement au dessus dans l'AOC-poset

Les concepts immédiatement en dessous dans l'AOC-poset

Les concepts minimaux qui sont en relation avec le point de départ

Navigation conceptuelle dans RCA

Objectif

Point de départ : un concept.
On veut :

- sa couverture supérieure ;
- sa couverture inférieure ;
- sa couverture relationnelle.

On veut également étendre le concept de départ avec les nouveaux attributs relationnels découverts.

Navigation conceptuelle dans RCA

L'exploration ne s'arrête pas à la première étape

On souhaite pouvoir faire plusieurs pas.

- Mettre à jour le contexte de départ avec des attributs relationnels
- Ne pas calculer le treillis du contexte cible en entier.

Navigation conceptuelle dans RCA

L'exploration ne s'arrête pas à la première étape

On souhaite pouvoir faire plusieurs pas.

- Mettre à jour le contexte de départ avec des attributs relationnels
- Ne pas calculer le treillis du contexte cible en entier.

Problèmes

- Définir où on veut aller ;
- information partielle : les opérateurs de dérivation ne fonctionnent plus de la même manière ;
- prendre en compte les relations cycliques.

Où veut-on aller ?

Définissons notre **strategie**

C notre point de départ, provient du contexte K_i. Une strategie est un ensemble de paires (r_{ij}, ρ) où

- r_{ij} est une relation sortant de notre contexte de départ ;
- ρ est le quantifieur choisi

La strategie nous permet, à chaque étape, de savoir quelles relations on veut prendre en compte et avec quel quantifieur.

Où veut-on aller ?

Définissons notre **strategie**

C notre point de départ, provient du contexte K_i. Une strategie est un ensemble de paires (r_{ij}, ρ) où

- r_{ij} est une relation sortant de notre contexte de départ ;
- ρ est le quantifieur choisi

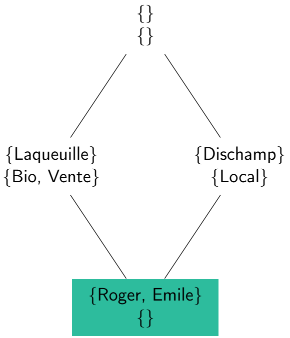
La strategie nous permet, à chaque étape, de savoir quelles relations on veut prendre en compte et avec quel quantifieur.

- Définir où on veut aller ✓

Nouveaux opérateurs de dérivation

Pourquoi les anciens ne fonctionnent pas ?

On ne calcule pas tous les concepts du contexte cible → on n'a pas tous les attributs relationnels du contexte de départ.



Nouveaux opérateurs de dérivation

Pourquoi les anciens ne fonctionnent pas ?

On ne calcule pas tous les concepts du contexte cible $\rightarrow$ on n'a pas tous les attributs relationnels du contexte de départ.

- Certains attributs ne sont connus qu'implicitement.
- L'intersection d'ensemble d'attributs dépend du quantifieur.

Besoin d'un algorithme pour calculer l'intersection d'ensembles d'attributs.

Nouveaux opérateurs de dérivation

Pourquoi les anciens ne fonctionnent pas ?

On ne calcule pas tous les concepts du contexte cible $\rightarrow$ on n'a pas tous les attributs relationnels du contexte de départ.

- intersection d'ensembles d'attributs ;
- intension d'un ensemble d'objets ;
- extension d'un ensemble d'attributs relationnels.

Nouveaux opérateurs de dérivation $\checkmark$

On étend partiellement les contextes

L'exploration ne s'arrête pas à la première étape

- Concepts introducteurs du contexte cible ;
- ajout de nouveaux attributs relationnels dans le contexte de départ.

On étend partiellement les contextes

L'exploration ne s'arrête pas à la première étape

- Concepts introducteurs du contexte cible ;
- ajout de nouveaux attributs relationnels dans le contexte de départ.

Permet une exploration étape par étape (avec une profondeur bornée).

Limitations de notre approche

Relations cycliques

Pas prises en compte.

Limitations de notre approche

Relations cycliques

Pas prises en compte.

Inconvénient

Toutes les configurations ne sont pas atteignables (en particulier le point fixe).

Limitations de notre approche

Relations cycliques

Pas prises en compte.

Inconvénient

Toutes les configurations ne sont pas atteignables (en particulier le point fixe). C'est grave ? Pas vraiment, dans la mesure où on veut juste explorer une partie de la famille de treillis.

Limitations de notre approche

Relations cycliques

Pas prises en compte.

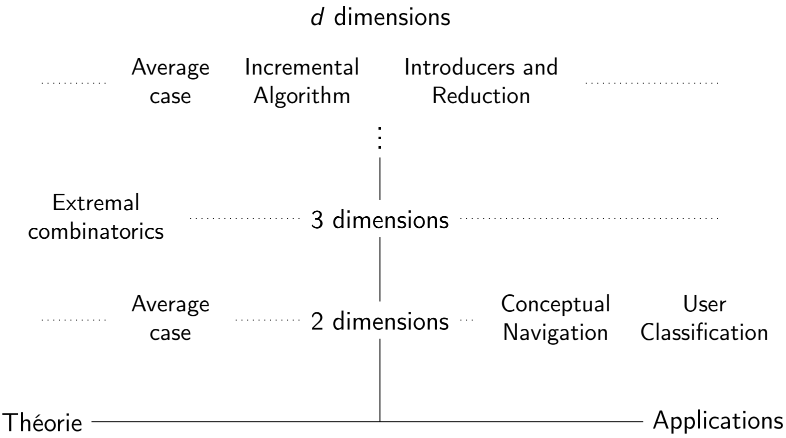
Inconvénient

Toutes les configurations ne sont pas atteignables (en particulier le point fixe). C'est grave ? Pas vraiment, dans la mesure où on veut juste explorer une partie de la famille de treillis.

Questions

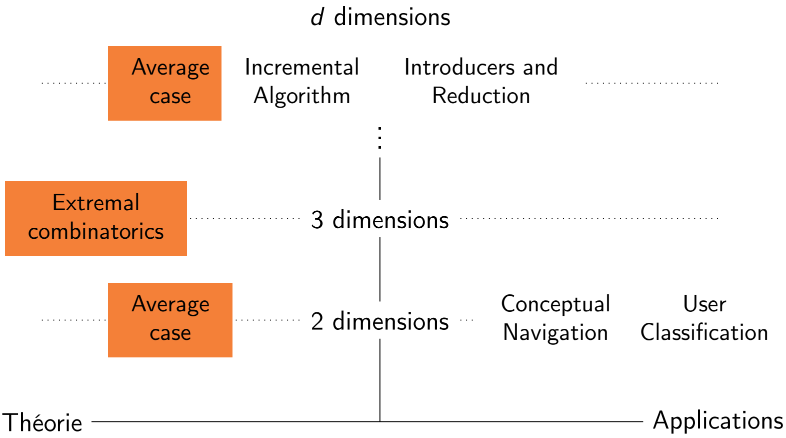
- Peut-on atteindre tous les concepts de la RCF depuis n'importe quel concept de départ ?
- Ce processus converge-t-il de la même manière que RCA ?

Conclusion

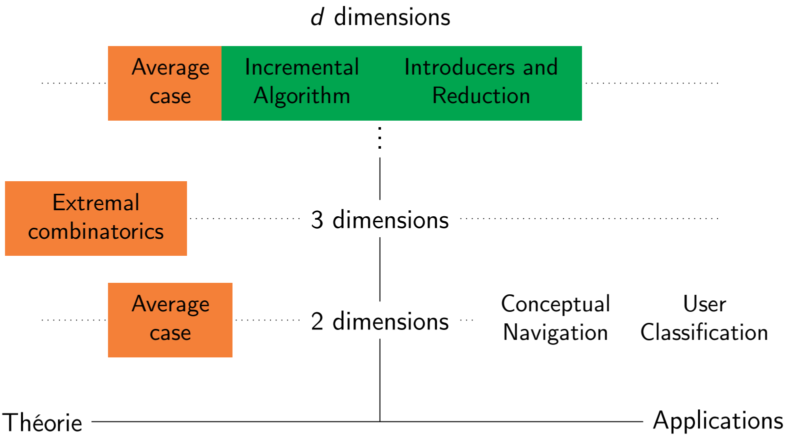


d dimensions
Average case
Incremental Algorithm
Introducers and Reduction
Extremal combinatorics
3 dimensions
Average case
2 dimensions
Conceptual Navigation
User Classification
Théorie
Applications

Conclusion

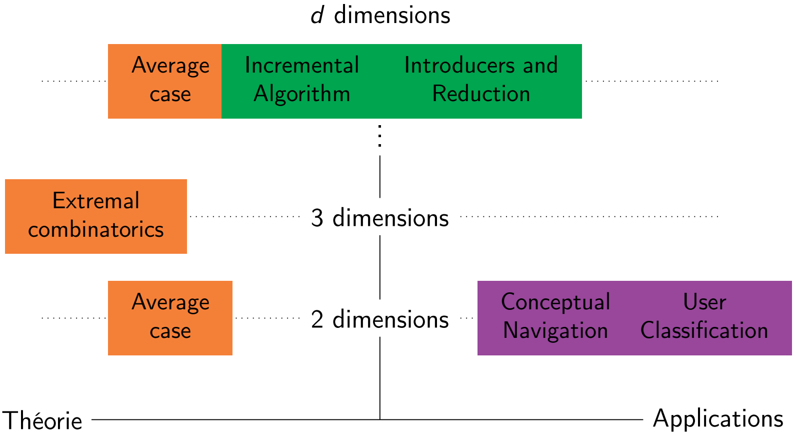

d dimensions
Average case
Incremental Algorithm
Introducers and Reduction
Extremal combinatorics
3 dimensions
Average case
2 dimensions
Conceptual Navigation
User Classification
Théorie
Applications

Conclusion



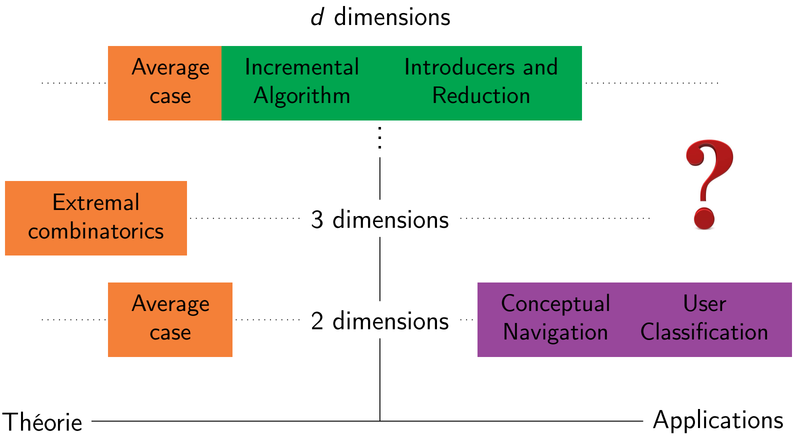
d dimensions
Average case
Incremental Algorithm
Introducers and Reduction
Extremal combinatorics
3 dimensions
Average case
2 dimensions
Conceptual Navigation
User Classification
Théorie
Applications

Conclusion



d dimensions
Average case
Incremental Algorithm
Introducers and Reduction
Extremal combinatorics
3 dimensions
Average case
2 dimensions
Conceptual Navigation
User Classification
Théorie
Applications

Perspectives



Merci de votre attention.

Productions

- *A Closed Set Based Learning Classifier for Implicit Authentication in Web Browsing*, avec D. Dia, F. Labernia, Y. Loiseau et O. Raynaud, *Discrete Applied Mathematics*. À paraître.
- *Average Size of implicational Bases*, avec A. Bazin, *CLA* 2018.
- *Bounding the Number of Minimal Transversals in Tripartite 3-Uniform Hypergraphs*, avec A. Bazin, L. Beaudou, et K. Khoshkhah, soumis.
- *On-Demand Relational Concept Analysis*, avec A. Bazin, J. Carbonnel et M. Huchard, document de travail.

Productions

- *On-Demand Generation of AOC-Posets : Reducing the Complexity of Conceptual Navigation*, avec A. Bazin et J. Carbonnel, *ISMIS* 2017.
- *Introducer Concepts in n-Dimensional Contexts*, avec A. Bazin, document de travail.
- *An Incremental Algorithm for Computing n-Dimensional Concepts*, avec A. Bazin et O. Raynaud, soumis.
- *A Tool for Classification of Sequential Data*, G. Kahn, Y. Loiseau et O. Raynaud, *FCA4AI@ECAI* 2016.

Productions

- *Bisplit graphs satisfy the Chen-Chvátal conjecture* avec L. Beaudou et M. Rosenfeld, soumis.
- *Encoding Partial Orders Through Modular Decomposition*, avec L. Beaudou, K. Ghazi, O. Raynaud et E. Thierry, *Journal of Computational Science* après *COMPSE* 2016.