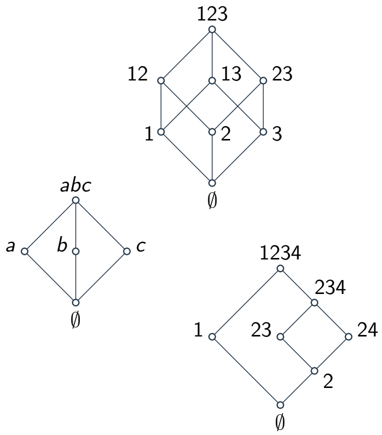
Translating between the representations of a ranked convex geometry

Oscar Defrain, Lhouari Nourine and *Simon Vilmin*.

LIMOS, UCA

WEPA - October, 2019

Underlying Structure



12 shortcut for $\{1, 2\}$.

- X set of elements, 2^X all subsets of X,
- if $\mathfrak{F} \subseteq 2^X$, $(X, \mathfrak{F})$ is a *set system*.

Definition - closure system

The pair $(X, \mathfrak{F})$ is a *closure system* if:

- $X \in \mathfrak{F}$,
- $F_1, F_2 \in \mathfrak{F} \implies F_1 \cap F_2 \in \mathfrak{F}$.

- $F \in \mathfrak{F}$ is called *closed*,
- $\phi(A) = min_{\subseteq}\{F \in \mathfrak{F} \mid A \subseteq F\}$ is the *closure* of A,
- uses in Computer Science: databases, Horn logic, data mining, . . .

General Problem

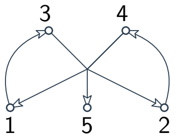
- *Problem*: a system $(X, \mathcal{F})$ requires *large amount of space*.
- *Solution*: use *implicit representations* of $(X, \mathcal{F})$:
  - *rules* allowing or not sets in $\mathcal{F}$,
  - a *minimum generating subset* of $\mathcal{F}$.

Implications

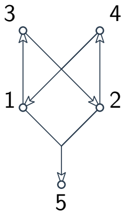
Definition - Implicational base

An *implicational base* is a pair (X, Σ) with :

- groundset X,
- Σ a set of *implications* $A \to B$, $A, B \subseteq \mathrm{X}$.



$\Sigma_1 = \{1 \to 3, 2 \to 4, 34 \to 125\}$.



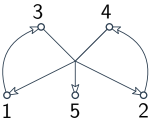
$\Sigma_2 = \{12 \to 5, 1 \to 3, 2 \to 4, 34 \to 12\}$.

- Intuition behind implication $A \to B$: *"we cannot have A without B"*,
- $F \subseteq \mathrm{X}$ *satisfies* $A \to B$ if $A \subseteq F$, then $B \subseteq F$,
- $\mathcal{F}_\Sigma = \{F \subseteq \mathrm{X} \mid X \text{ satisfies } \Sigma\}$,
- also known as : *directed hypergraph*, functional dependencies, Horn clauses, . . .

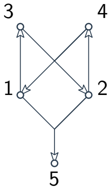
Implications

Theorem (folklore)

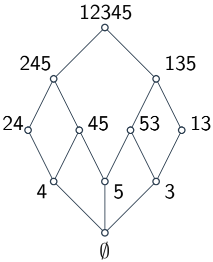
- $(X, \mathcal{F}_\Sigma)$ is a closure system,
- every closure system arises from some implicational base (X, Σ).



Σ_1



Σ_2



$\mathcal{F}_{\Sigma_1} = \mathcal{F}_{\Sigma_2}$

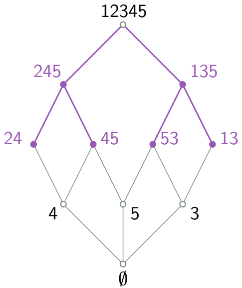
Beware ! The correspondence is *not* one-to-one !

Generating Subset

Definition - irreducible element

For a closure system $(X, \mathcal{F})$, $M \in \mathcal{F}$ is a *(meet-)irreducible element* if for all $F_1, F_2 \in \mathcal{F}$:

- $F_1 \cap F_2 = M \implies M = F_1$ or $M = F_2$.



- $\mathcal{M}$ irreducible elements of $\mathcal{F}$,
- $M \in \mathcal{M}$ iff it has a *unique* up successor in the diagram,
- any $F \in \mathcal{F} \setminus \mathcal{M}$ is the intersection of some irreducible $\implies \mathcal{M}$ *generates* $\mathcal{F}$,
- known as: *characteristic models*, MAX-sets, copoints, ...

Enumeration problems

Problem - Computing Characteristic Models (CCM)

Input: an implicational base (X, Σ).

Output: the set $\mathcal{M}$ of irreducible elements of $(X, \mathcal{F}_\Sigma)$.

Problem - Structure Identification (SID)

Input: the set $\mathcal{M}$ of irreducible elements of $(X, \mathcal{F})$.

Output: a *minimum* implicational base (X, Σ) such that $\mathcal{F}_\Sigma = \mathcal{F}$.

Known Results

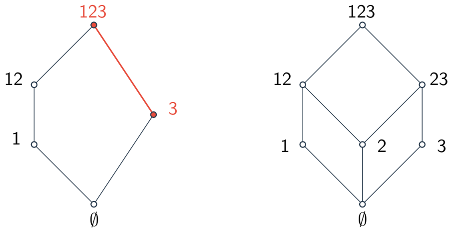
- Still open problems,
- but *harder than* minimal transversals enumeration [Khardon, 1995],
- algorithms for classes of closure systems [Beaudou et al., 2017],
- algorithms for particular implicational bases [Korte et al., 2012], [Adaricheva, Nation, 2017].

Convex Geometries

Definition - Convex Geometry (CG)

A closure system $(X, \mathcal{F})$ is a *convex geometry* (CG) if it is strongly accessible, i.e.:

- $\forall F \in \mathcal{F}, \exists x \in X \setminus F$, s.t. $F \cup \{x\} \in \mathcal{F}$.



- strong accessibility $=$ choose elements of X one by one,
- seen in: learning spaces, antimatroids, social choice operators, ...

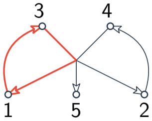
Acyclic convex geometries

Definition - implication graph

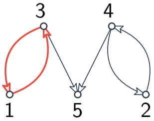
The *implication graph* $G_\Sigma = (X, E)$ of an implicational base (X, Σ) is a *directed* graph where $(a, b) \in E$ if there is $A \rightarrow B \in \Sigma$ such that $a \in A, b \in B$.

Definition - Acyclic implicational base

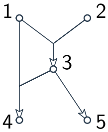
An implicational base (X, Σ) is *acyclic* if G_Σ does not have cycles.



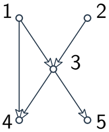
Σ_1



G_{Σ_1}



Σ_2



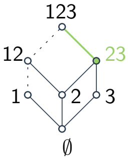
G_{Σ_2}

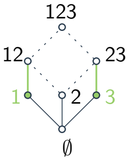
Theorem (Wild, 94)

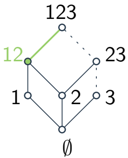
(X, Σ) acyclic $\implies$ $(X, \mathcal{F}_\Sigma)$ convex geometry.

Why Acyclicity ?

- Structural properties of $\mathcal{M}$:
  - for any $x \in X$, $max_{\subseteq}\{M \in \mathcal{F} \mid x \notin M\} \subseteq \mathcal{M}$,
  - in convex geometries, *partition* of $\mathcal{M}$!
  - suggests enumeration for each $x \in X$.
- Hardness bounds:
  - still *harder than* minimal transversals enumeration,
  - even *harder than* dualization in distributive closure systems [Defrain, Nourine, V., 2019],
  - is there an *"easy"* subclass in between ?

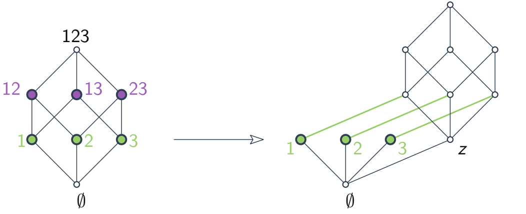




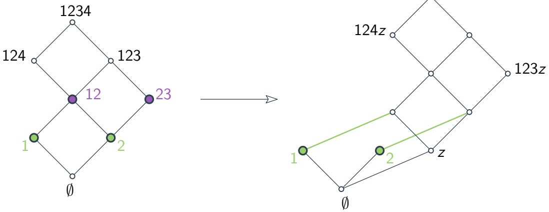


Dualization Parenthesis

- Minimal transversals : *no* constraints ($\Sigma = \emptyset$),
- hypergraph becomes implications,
- *acyclic* implicational system.



$\Sigma = \emptyset$ $\qquad$ $\Sigma = \{12 \to z, 23 \to z, 13 \to z\}$



$\Sigma = \{4 \to 12, 3 \to 2\}$ $\qquad$ $\Sigma = \{4 \to 12, 3 \to 2\} \cup \{12 \to z, 3 \to z\}$

- generalization : constraints ($\Sigma \neq \emptyset$),
- *unitary* implications,
- remains *acyclic*.

Ranked Acyclic System

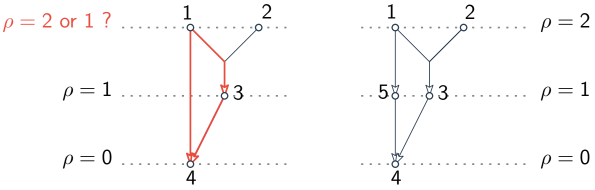
Definition - rank function

A *rank function* on an implicational base (X, Σ) is a function $\rho : \mathsf{X} \longrightarrow \mathbb{N}$ such that:

$$A \rightarrow B \in \Sigma \implies \rho(a) = \rho(b) + 1, \text{ for } a \in A, b \in B.$$

Definition - ranked implicational base, ranked convex geometry

An implicational base (X, Σ) is *ranked* if it admits a rank function. The system $(\mathsf{X}, \mathcal{F}_\Sigma)$ is then called a *ranked convex geometry*.



Main algorithmic results

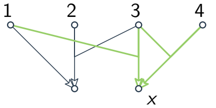
Theorem (Defrain, Nourine, V., 2019)

Let (X, Σ) be a ranked implicational base. There exists an output-quasi polynomial time algorithm for CCM with input (X, Σ).

Theorem (Defrain, Nourine, V., 2019)

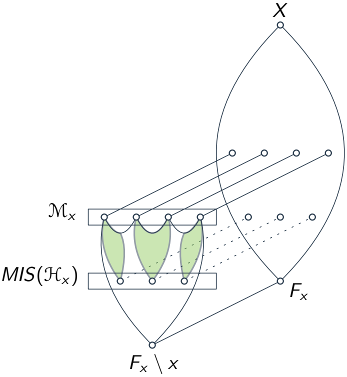
Let $\mathcal{M}$ be the set of irreducible elements of a ranked convex geometry. There exists an output-quasi polynomial time algorithm for SID with input $\mathcal{M}$.

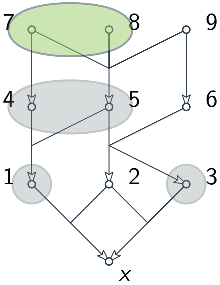
CCM Algorithm



$\mathcal{H}_x = \{13 \to x, 34 \to x\}$

- (X, Σ) a ranked base.
- Structural insights, $x \in X$:
  - hypergraph $\mathcal{H}_x = \{A \to x \mid A \to x \in \Sigma\}$,
  - $\mathcal{M}_x$ irreducible associated to x,
  - $\mathcal{M}_x$ *partitionned* by $MIS(\mathcal{H}_x)$.
- Algorithm outline:
  - *recursive* application of minimal transversals enumeration, rank by rank,
  - height of the recursive tree $\leq \rho$,
  - apply for all $x \in X$.





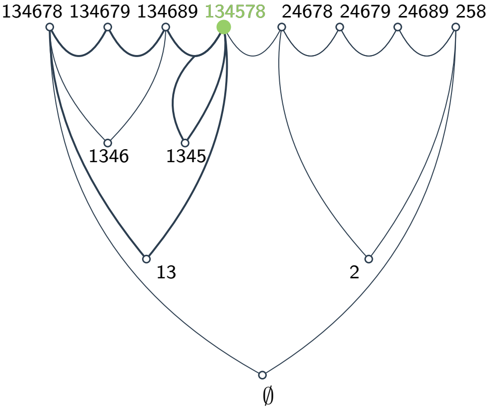
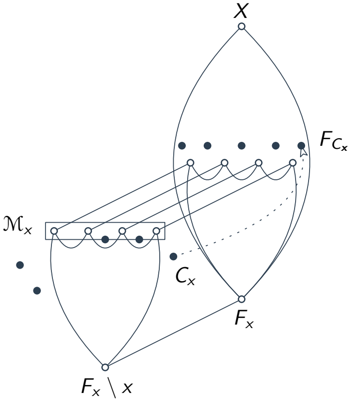


Figure: $\Sigma = \{7 \to 4, 789 \to 5, 9 \to 6, 45 \to 1, 56 \to 23, 12 \to x, 23 \to x\}$

SID Algorithm

- $\mathcal{M}$ of a ranked CG given.
- Structural insights:
  - C_x generates F_{C_x},
  - for any $M_x \in \mathcal{M}_x$, $F_{C_x} \nsubseteq M_x \cup \{x\}$,
  - F_{C_x} min for this property (*dualization*),
  - $\Sigma = \{C_x \to x \mid x \in X\}$ is minimum.
- Algorithm outline, for $x \in X$:
  - identify elements of C_x using $\mathcal{M}_x$,
  - hypergraph based on $X \setminus M_x$ for $M_x \in \mathcal{M}_x$,
  - minimal transversal enumeration to find C_x's.



Conclusion

- What we discussed so far:
  - closure systems, their representations,
  - translating algorithms for *ranked* convex geometries $(\mathrm{SID}, \mathrm{CCM})$.
- Further questions:
  - *recognition* of a ranked CG from $\mathcal{M}$,
  - apply the algorithm to a *broader class* of CG.

Thank you for your attention !

References

- K. Adaricheva and, J.B. Nation.
  Discovery of the D-basis in binary tables based on hypergraph dualization.
  Theoretical Computer Science, 658 :307-315, 2017.
- L. Beaudou, A. Mary, and L. Nourine.
  Algorithms for *k*-meet-semidistributive lattices.
  Theoretical Computer Science, 658 :391-398, 2017.
- O. Defrain, L. Nourine, and S. Vilmin.
  Translating between the representations of a ranked convex geometry.
  Arxiv, 1907.09433, 2019
- R. Khardon.
  Translating between Horn Representations and their Characteristic Models.
  Journal of Artificial Intelligence Research, 3 :349-372, 1995.
- B. Korte, L. Lovasz and R. Schrader
  Greedoids
  Springer Science & Business Media, 2012.
- M. Wild,
  A Theory of Finite Closure Spaces Based on Implications
  Technische Hochshule Darmstadt