Dihypergraph decomposition: application to closure system representations

8th FCA4AI Workshop

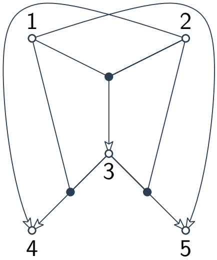
Lhouari Nourine and *Simon Vilmin*.

LIMOS, UCA

August 2020

Implications, Dihypergraphs

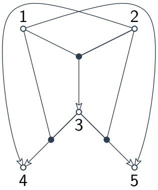
- Set V of vertices (attributes)
- Dependencies in V: implications $B \rightarrow h$, $B \subseteq \mathsf{V}$, $h \in \mathsf{V}$
- Represented by a *dihypergraph* $\mathcal{H} = (\mathsf{V}, \mathcal{E})$, the arc (B, h) models $B \rightarrow h$

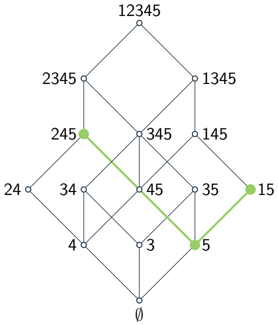


$(12, 3)$
$(1, 5)$
$(2, 4)$
$(13, 4)$
$(23, 5)$

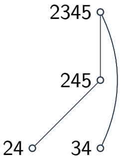
Dihypergraphs, Closure systems

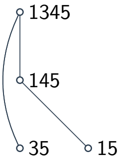
- F *models* (B, h) if $B \subseteq F \implies h \in F$
- F *closed* in $\mathcal{H}$ if F models $\mathcal{E}$ (forward chaining)
- $\mathcal{F} = \{F \subseteq V \mid F \text{ closed in } \mathcal{H}\}$ is a *closure system*:
  - ▷ $V \in \mathcal{F}$
  - ▷ $F_1, F_2 \in \mathcal{F} \implies F_1 \cap F_2 \in \mathcal{F}$





Closure systems, Meet-irreducible

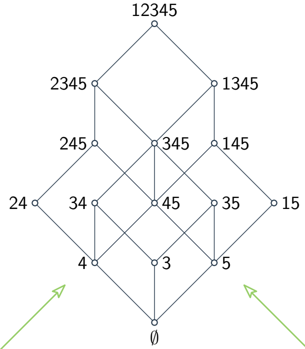




	1	2	3	4	5
24		×		×	
34			×	×	
35			×		×
15	×				×
245		×		×	×
145	×			×	×
2345		×	×	×	×
1345	×		×	×	×

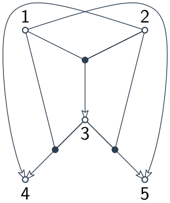
- $\emptyset$ obtained by intersection, brings no informationsame for $4, 3, 5, 45, 345, 1234524$ cannot be obtained, it is meet-irreducible Meet-irreducible $\mathcal{M} \equiv$ reduced context

Problem



poly

poly



???

	1	2	3	4	5
24		×		×	
34			×	×	
35			×		×
15	×				×
245		×		×	×
145	×			×	×
2345		×	×	×	×
1345	×		×	×	×

Problem

Problem - Enumerating Meet-Irreducible

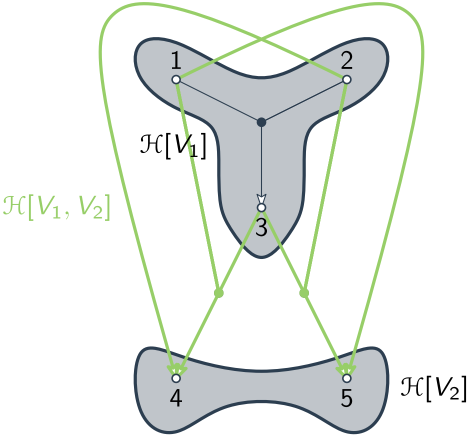
- **Input:** a dihypergraph $\mathcal{H} = (V, \mathcal{E})$.
- **Output:** the set $\mathcal{M}$ of meet-irreducible elements of $\mathcal{F}$.

- survey in [Bertet et al., 2018]
- Negative side:
  - *harder than* hypergraph dualization [Khardon, 1995]
  - pseudo-intent recognition *coNP-C* [Babin, Kuznetsov, 2013]
- Positive side:
  - generic algorithms [Mannila, Räihä, 1992]
  - classes of closure systems [Beaudou et al., 2017, Defrain et. al., 2019]

Strategy

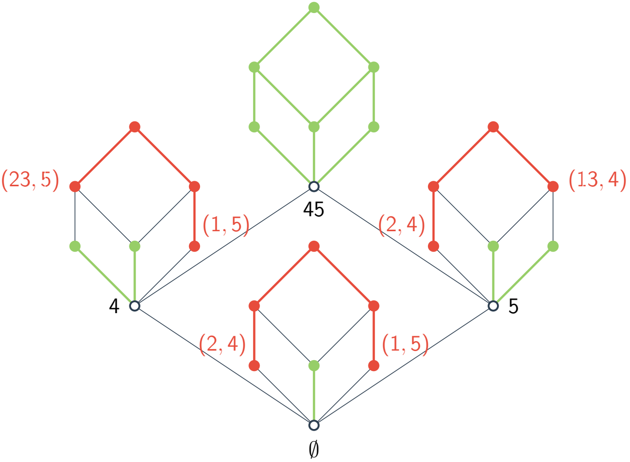
- *Acyclic split* of $\mathcal{H}$:
  - bipartition $(\mathsf{V}_1, \mathsf{V}_2)$ of V s.t. any arc (B, h) is either in V_1, in V_2 or $B \subseteq \mathsf{V}_1$ and $h \in \mathsf{V}_2$
  - partitions $\mathcal{H}$ into $\mathcal{H}[\mathsf{V}_1]$, $\mathcal{H}[\mathsf{V}_2]$ and a *bipartite dihypergraph* $\mathcal{H}[\mathsf{V}_1, \mathsf{V}_2]$
- Application :
  - characterization of $\mathcal{M}$
  - recursive application to obtain hierarchical decomposition of $\mathcal{H}$

Acyclic split



Closure system construction

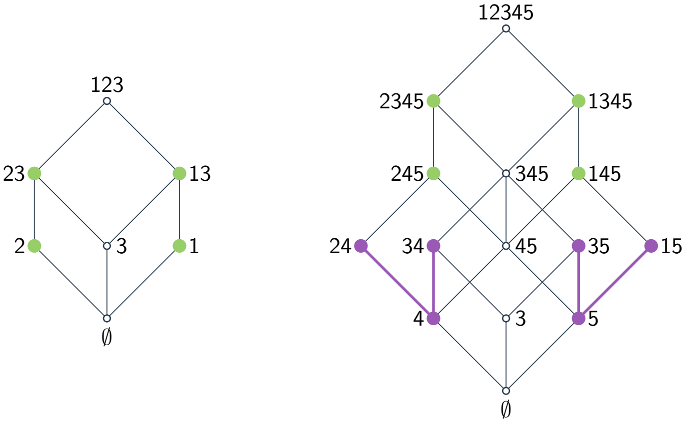
- First case: $\mathcal{H}[V_1, V_2]$ has *no* arcsSecond case: $\mathcal{H}[V_1, V_2]$ has arcs
- $F_2 \in \mathcal{F}_2$ combined with any $F_1 \in \mathcal{F}_1$copies of $\mathcal{F}_1$ on each $F_2 \in \mathcal{F}_2\mathcal{F}$ direct product of $\mathcal{F}_1, \mathcal{F}_2$*Extensions* of F_2 are controlled by $\mathcal{H}[V_1, V_2]$Increasing copies of *ideals* of $\mathcal{F}_1$ on $\mathcal{F}_2$



Formal sum-up

- (V_1, V_2) is an acyclic split of $\mathcal{H}$
- *trace* (projection) $\mathcal{F} : V_1 = \{F \cap V_1 \mid F \in \mathcal{F}\}$
- $\mathcal{F}$ is built by adding parts of $\mathcal{F}_1$ to each $F_2 \in \mathcal{F}_2$:
  - $\text{Ext}(F_2) = \{F \in \mathcal{F} \mid F \cap V_2 = F_2\}$, extensions of $F_2 \in \mathcal{F}_2$
  - $\text{Ext}(F_2)$ corresponds to an ideal of $\mathcal{F}_1$ controlled by $\mathcal{H}[V_1, V_2]$
  - $F_2 \subseteq F_2'$ implies $\text{Ext}(F_2) : V_1 \subseteq \text{Ext}(F_2') : V_1$
  - $F_2' \succ F_2$ implies that extensions of F_2' cover extensions of F_2
- Note: $\text{Ext}(V_2) : V_1$ is $\mathcal{F}_1$

Meet-irreducible identification



Meet-irreducible characterization

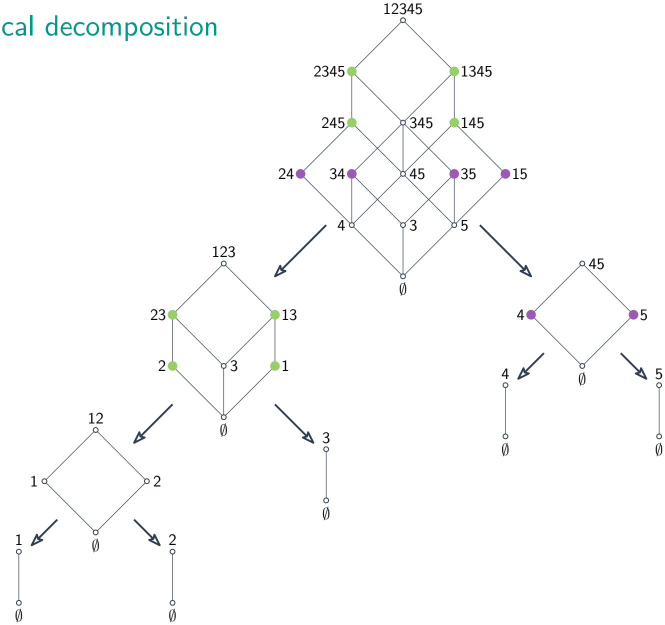
Theorem (Nourine, V., 2020+)

Let $\mathcal{H} = (\mathsf{V}, \mathcal{E})$ be a dihypergraph and $(\mathsf{V}_1, \mathsf{V}_2)$ an acyclic split of $\mathcal{H}$. Meet-irreducible elements of $\mathcal{M}$ of $\mathcal{F}$ are given by the following equality:

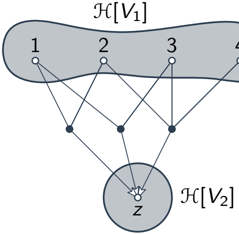
$$\mathcal{M} = \{M_1 \cup \mathsf{V}_2 \mid M_1 \in \mathcal{M}_1\} \cup \{F \in \max_{\subseteq}(\mathsf{Ext}(M_2)) \mid M_2 \in \mathcal{M}_2\}$$

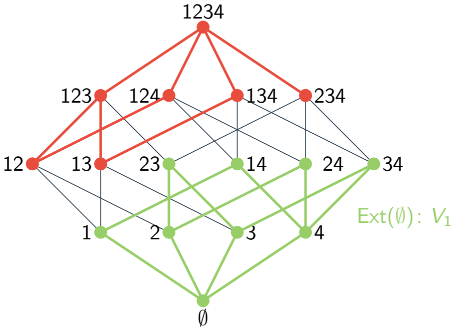
where $\mathcal{M}_1, \mathcal{M}_2$ are meet-irreducible elements of $\mathcal{H}[\mathsf{V}_1], \mathcal{H}[\mathsf{V}_2]$ respectively.

Hierarchical decomposition



Finding maximal extensions





- $\mathcal{H}[V_1, V_2]$ defines the antichain of minimum forbidden extensions Pick an hypergraph $\mathcal{H} = (V, \mathcal{E})$, turn each arc B into (B, z) Acyclic split (V, z) Extensions of $\emptyset$ are determined by $MIS(\mathcal{H})$

Finding maximal extensions

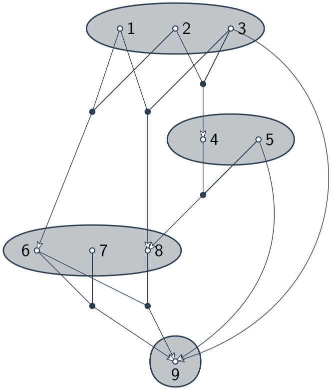
Problem - Finding Maximal Extensions (FME)

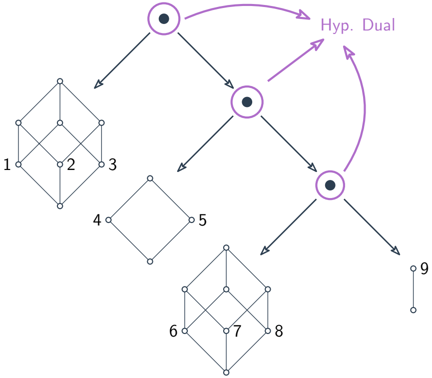
- **Input:** a dihypergraph $\mathcal{H} = (V, \mathcal{E})$ with acyclic split (V_1, V_2), $\mathcal{H}[V_1]$, $\mathcal{H}[V_2]$, $\mathcal{H}[V_1, V_2]$, $\mathcal{M}_1, \mathcal{M}_2$ and $F_2 \in \mathcal{F}_2$.
- **Output:** $\max_{\subseteq}(\mathrm{Ext}(F_2))$.

- Open in general
- if $\mathcal{H}[V_1]$ has no arcs, FME $\equiv$ hypergraph dualization!

$\Longrightarrow$ output quasi-polynomial time algorithm [Fredman, Khachiyan, 1996]

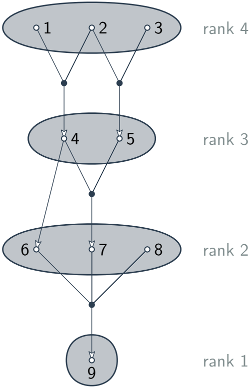
New tractable cases





Main idea: connecting blocks in an acyclic way

Ranked convex geometries



- ranked convex geometry

Theorem [Defrain et. al., 2019]

Enumerating meet-irreducible elements $\mathcal{M}$ from $\mathcal{H}$ in ranked convex geometries can be done in output quasi-polynomial time.

Conclusion

- Problem:
  - enumeration of meet-irreducible elements of a dihypergraph (implications)
  - harder than hypergraph dualization in general
- Our contribution (Nourine, V., 2020+):
  - use of a decomposition operation, *acyclic split*
  - recursive characterization of $\mathcal{M}$
  - application to new tractable classes of closure systems, generalizing [Defrain et. al., 2019]
- Future works:
  - characterize/improve tractable cases
  - translation in the other way

Thank you for your attention!

References

- R. Khardon.
  Translating between Horn Representations and their Characteristic Models.
  Journal of Artificial Intelligence Research, 3 :349-372, 1995.
- O. Defrain, L. Nourine, S. Vilmin.
  Translating between the representations of a ranked convex geometry.
  arXiv:1907.09433, 2019.
- M. Babin, S. Kuznetsov.
  Computing premises of a minimal cover of functional dependencies is intractable.
  Discrete Applied Mathematics, 161 :742-749, 2013.
- L. Beaudou, A. Mary, and L. Nourine.
  Algorithms for k-meet-semidistributive lattices.
  Theoretical Computer Science, 658 :391-398, 2017.
- H. Mannila, K.-J. Räihä.
  The design of relational databases.
  Addison-Wesley Longman Publishing Co., Inc., 1992.
- M. Fredman, L. Khachiyan.
  On the complexity of dualization of monotone disjunctive normal forms.
  Journal of Algorithms, 21 :618-628, 1996.
- K. Bertet, C. Demko, J.-F. Viaud, and C. Guérin.
  Lattices, closures systems and implication bases: A survey of structural aspects and algorithms.
  Theoretical Computer Science 743 :93-109, 2018.