

Algorithms on closure systems and their representations

Simon Vilmin

LIMOS, CNRS, Université Clermont Auvergne

December 13, 2021

Committee			
Kira	ADARICHEVA	Reviewer	
Karell	BERTET	Reviewer	
Sergeï	KUZNETSOV	Reviewer	
Arnaud	MARY	Examiner	
Jean-Marc	PETIT	Examiner	
Lhouari	NOURINE	Advisor	



Funded by the **ProFan Project**

- Initial motivation: *Knowledge Space Theory* [Doignon, Falmagne, 1985].
- Some questions of an automated test:
 1. Graphically solve $4x^2 - 3x + 2 = 0$.
 2. Figure out $\frac{\sqrt{4} \times \sqrt{9}}{3} - \frac{6 \times 7}{\sqrt{144}}$.
 3. Compute the discriminant of $3x^2 - x + 8$.
 4. Study the polynomial $7x^2 + 11x - 5$.
- Each question corresponds to a *problem* or *item*:
 1. Graphical resolution.
 2. Arithmetic.
 3. Formula of discriminant.
 4. Study of a 2nd order polynomial.

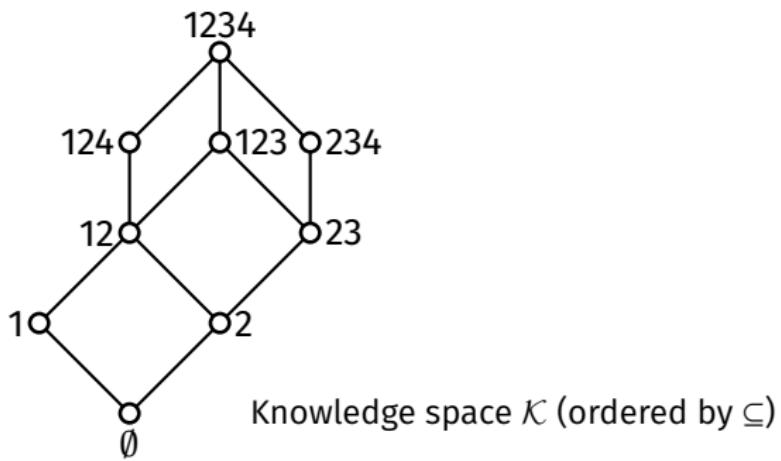
Introduction ▷ Time for results!

	1	2	3	4
Wolf	x			
Lil		x	x	
Lazuli		x	x	x
Folavril	x	x		x
Dupont		x		

- Some students took the test!
- Lazuli *masters* item 3.
- {2,3} is the *knowledge state* of Lil.

Introduction ▷ Knowledge spaces

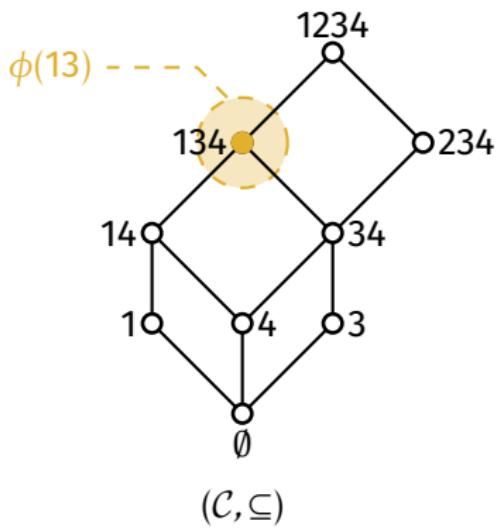
- *Knowledge space \mathcal{K} over a (finite) collection of items V :*



Definition ▷ Closure system

Closure system $\mathcal{C} \subseteq 2^V$ over V :

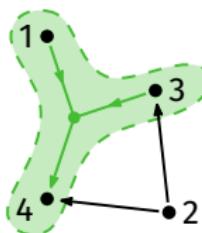
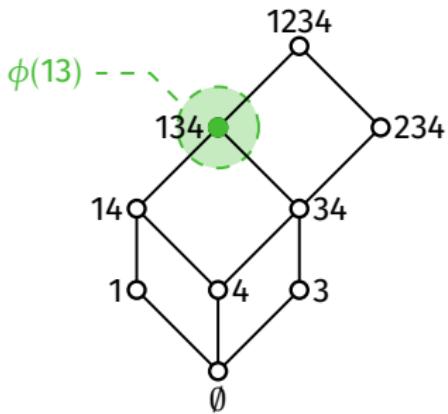
- Contains V .
- Closed by intersection: $C_1, C_2 \in \mathcal{C}$ entails $C_1 \cap C_2 \in \mathcal{C}$.



- Sets in \mathcal{C} are *closed sets*.
- (\mathcal{C}, \subseteq) is a (*closure*) lattice.
- Induces a *closure operator* ϕ :
 - $\phi(X)$: minimal closed set including X .
- Closure system = *complement* of Knowledge space!
- \mathcal{C} standard: $\phi(v) \setminus \{v\} \in \mathcal{C}$ for each $v \in V$.

- Closure systems are *ubiquitous* ...
 - *Knowledge Space Theory (KST)*,
 - *Formal Concept Analysis (FCA)*,
 - *Propositional logic*,
 - *Argumentation theory*,
 - *Databases*,
 - ...
- ... but they have *HUGE* size ...
 - If V has n elements, \mathcal{C} can have 2^n closed sets!
- ... and can be *hard to understand*:
 - In KST: asking teachers to provide raw knowledge states is impractical.
- We need *implicit representations!*

- *Implication*: expression $A \rightarrow B$, where $A, B \subseteq V$.
- *Implicational base*: set Σ of implications.
- “If the students fail the items in A , they will fail the items in B ”.
- Σ represents a *unique* closure system \mathcal{C} .
- \mathcal{C} can be represented by *several (equivalent)* Σ .



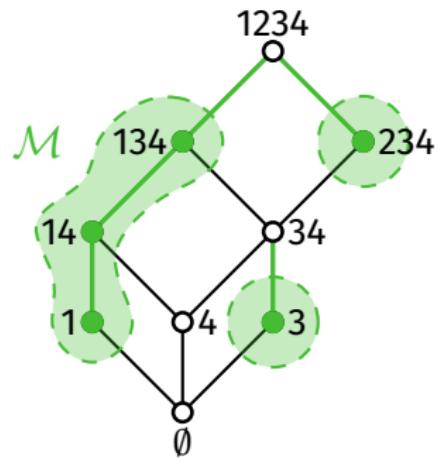
$$\Sigma = \{13 \rightarrow 4, 2 \rightarrow 34\}$$

Definition ▷ Meet-irreducible elements

Closure system \mathcal{C} over V :

- $M \in \mathcal{C} \setminus \{V\}$ *meet-irreducible* if $M = C_1 \cap C_2$ implies $M = C_1$ or $M = C_2$, $C_1, C_2 \in \mathcal{C}$.
- \mathcal{M} collection of all meet-irreducible elements of \mathcal{C} .

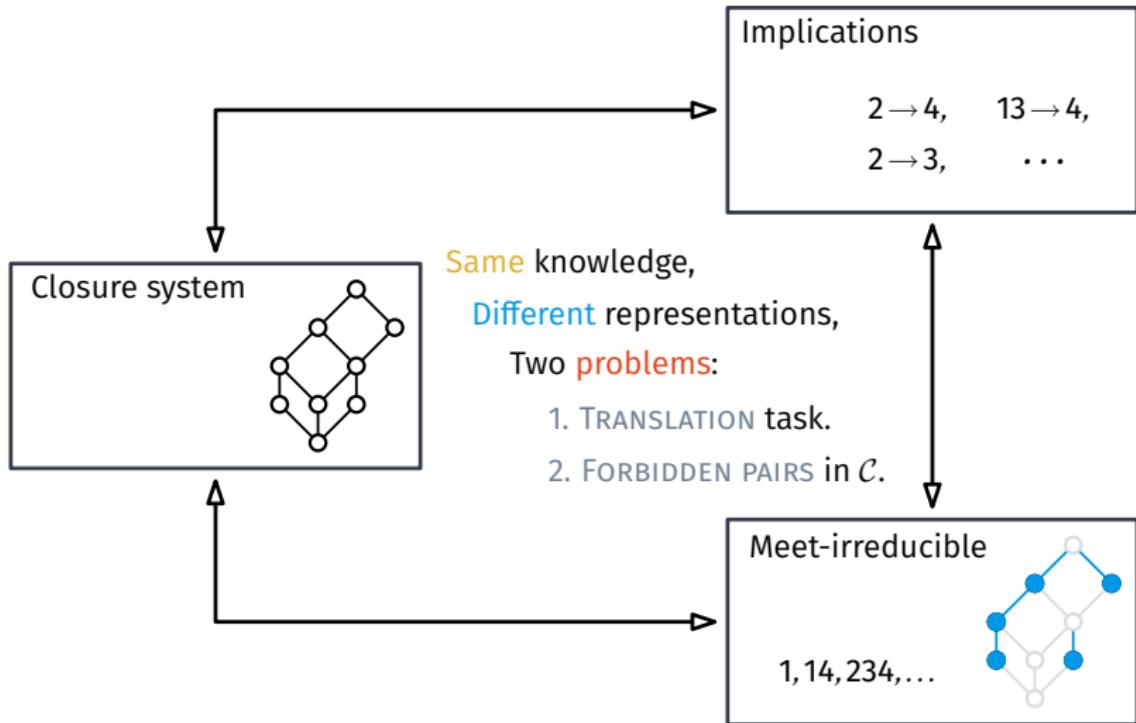
- \mathcal{C} *fully recovered* from \mathcal{M} by taking intersections.
- \mathcal{M} is the “*core*” of \mathcal{C} .
- $M \in \mathcal{M}$ iff unique *cover*.



Introduction ▷ Pros and cons

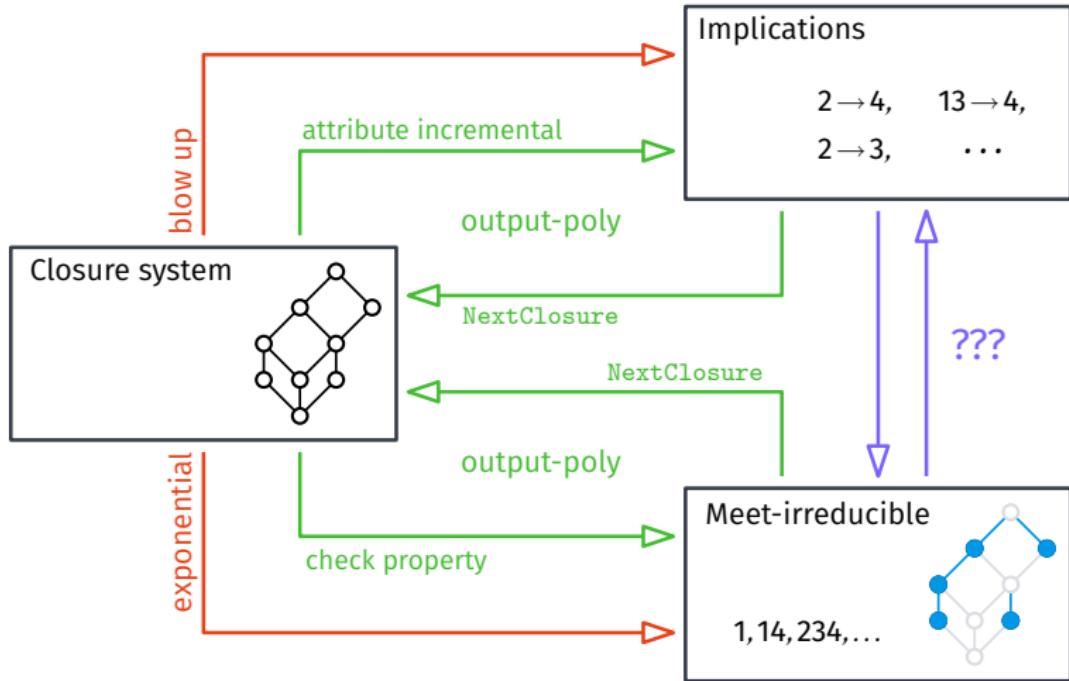
Question	Σ	\mathcal{M}	\mathcal{C}
is v in a min. generator of u ?	✗	✓	✓
is P pseudo-closed?	✓	✗	✓
is \mathcal{C} join-semidistributive?	?	✓	✓
Relative size			
size of ... w.r.t. Σ	—	$\exp(\Sigma)$	$\exp(\Sigma)$
size of ... w.r.t. \mathcal{M}	$\exp(\mathcal{M})$	—	$\exp(\mathcal{M})$
size of ... w.r.t. \mathcal{C}	$\leq \mathcal{C} \times V $	$\leq \mathcal{C} $	—

✓ Polynomial
✗ NP-complete



First problem ► Translating between the representations
of a closure system.

Translation ▷ Travelling between the representations



Problem ▶ ENUM. MEET-IRR. ELEMENTS (CCM)

- *Input:* an implicational base Σ for a closure system \mathcal{C} over V .
 - *Output:* the meet-irreducible \mathcal{M} of \mathcal{C} .
-
- Surveys by [Bertet et al., 2018], [Wild, 2017].
 - Hardness results:
 - *Unknown* complexity.
 - Harder than *hypergraph dualization* (MISENUM), [Khordon, 1995].
 - Enumerating co-atoms is intractable (dualization), [Kavvadias et al., 2000].
 - Positive results:
 - General (exponential) algorithms [Mannila, Räihä, 1992], [Wild, 1995].
 - Tractable cases: meet-semidistributive, types of convex geometries [Beaudou et al., 2017], [Defrain, Nourine, V., 2021].

- Strategy:
 - Hierarchical decomposition of Σ .
 - Recursive construction of \mathcal{M} .

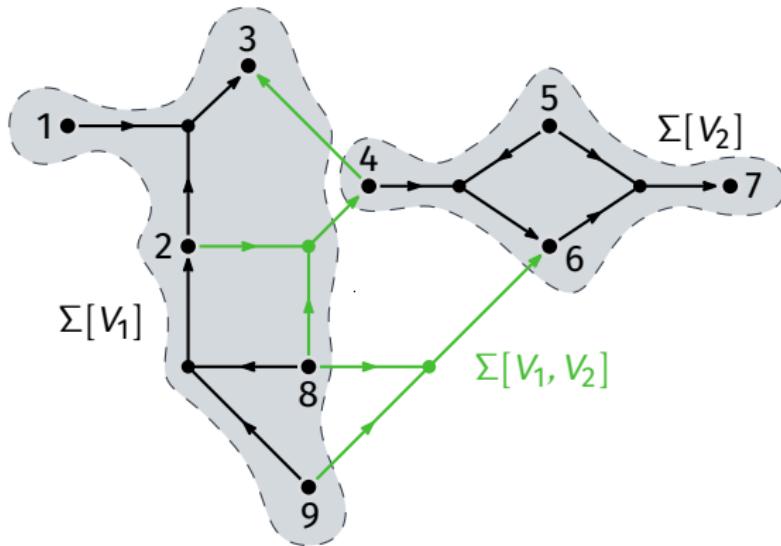
Definition ▷ Split

Σ implicational base over V :

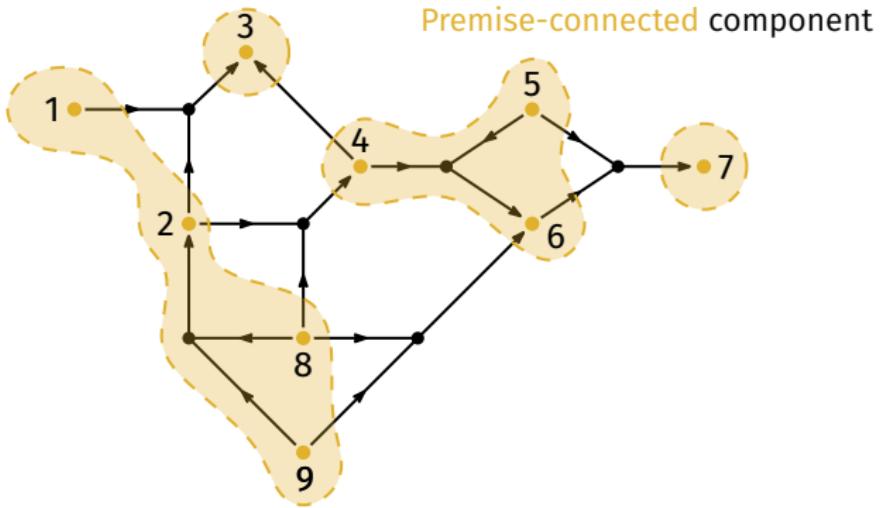
- *Split* of Σ : bipartition (V_1, V_2) of V such that $A \subseteq V_1$ or $A \subseteq V_2$ for every $A \rightarrow B \in \Sigma$.

- Split (V_1, V_2) partitions Σ :
 - $\Sigma[V_1]$ implications included in V_1 , with induced $\mathcal{C}_1, \mathcal{M}_1$.
 - $\Sigma[V_2]$ implications included in V_2 , with induced $\mathcal{C}_2, \mathcal{M}_2$.
 - $\Sigma[V_1, V_2]$ implications from V_1 to V_2 or from V_2 to V_1 .

Translation ▷ Split operation



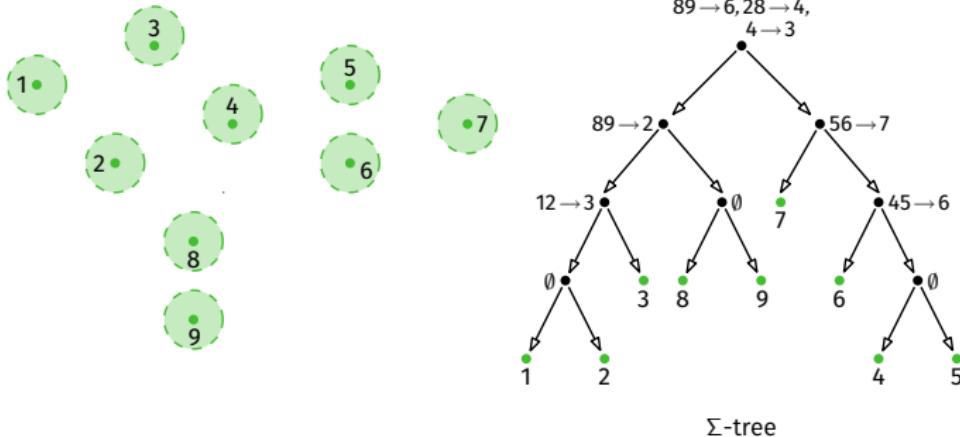
Translation ▶ Recognizing splits



Proposition ▶ Recognizing splits

Σ has a split (V_1, V_2) if and only if it is not premise-connected.

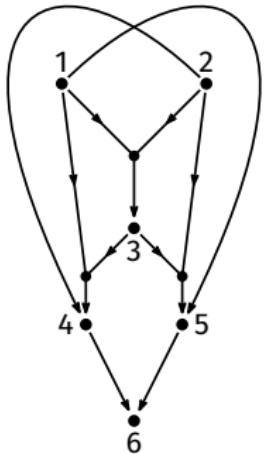
Translation ▷ Hierarchical Decomposition



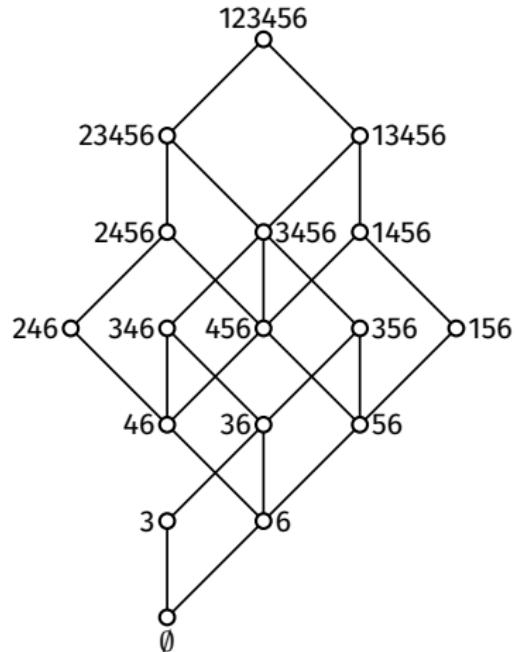
Theorem ▷ Nourine, V.

Let Σ be an implicational base over V . A Σ -tree can be computed in *polynomial time and space* in the size of Σ , if it exists.

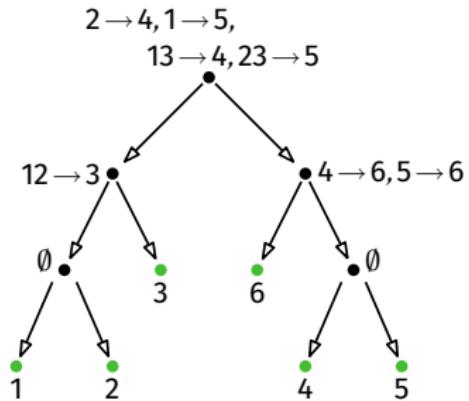
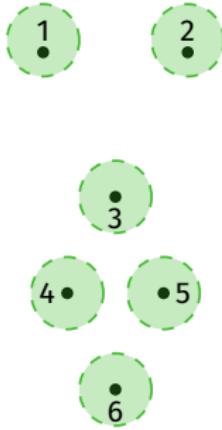
Translation ▷ Back to closure systems and CCM



$$\Sigma = \left\{ \begin{array}{l} 12 \rightarrow 3, \quad 1 \rightarrow 5, \quad 5 \rightarrow 6, \\ 13 \rightarrow 4, \quad 2 \rightarrow 4, \\ 23 \rightarrow 5, \quad 4 \rightarrow 6 \end{array} \right\}$$

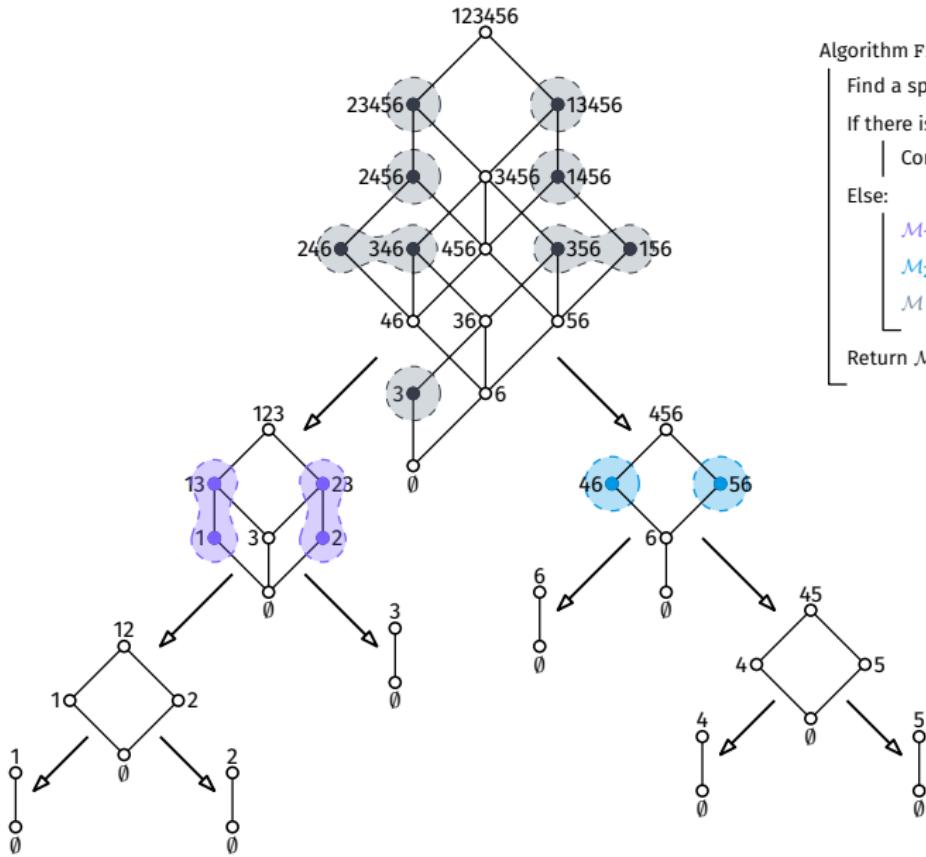


Translation ▷ Back to closure systems and CCM



- H-decomposition of Σ implies *H-decomposition* of \mathcal{C} .

Translation ▷ Back to closure systems and CCM



Algorithm `FindMeet(Σ, V)`

Find a split (V_1, V_2) of Σ

If there is none:

Compute \mathcal{M} with another algorithm

Else:

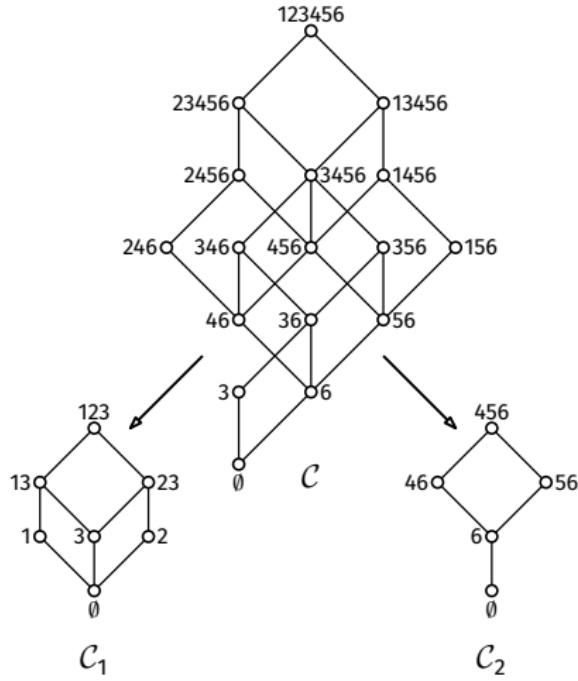
$\mathcal{M}_1 = \text{FindMeet}(\Sigma[V_1], V_1)$

$\mathcal{M}_2 = \text{FindMeet}(\Sigma[V_2], V_2)$

$\mathcal{M} = \text{ComputeMeet}(\mathcal{M}_1, \mathcal{M}_2, \Sigma)$

Return \mathcal{M}

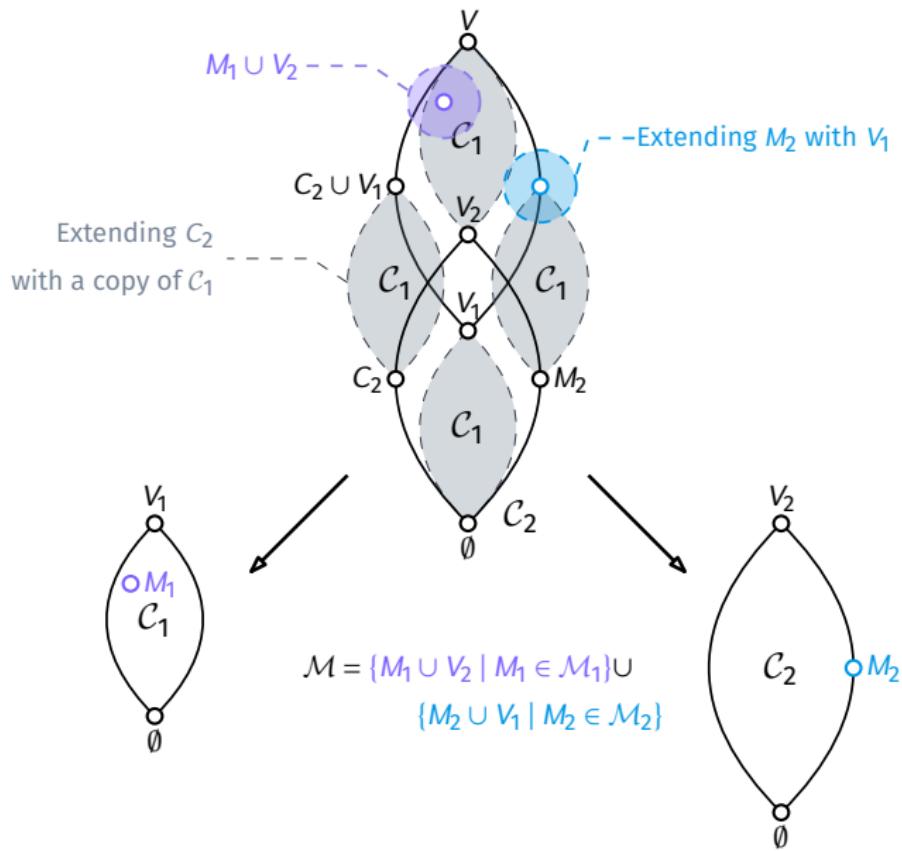
Translation ▷ Back to closure systems and CCM



Observation ▷ Closed sets

- $\mathcal{C} \subseteq \mathcal{C}_1 \times \mathcal{C}_2$.

Translation ▷ Constructing \mathcal{C}, \mathcal{M} with empty split

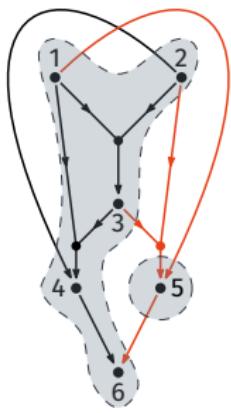


Translation ▷ Acyclic split

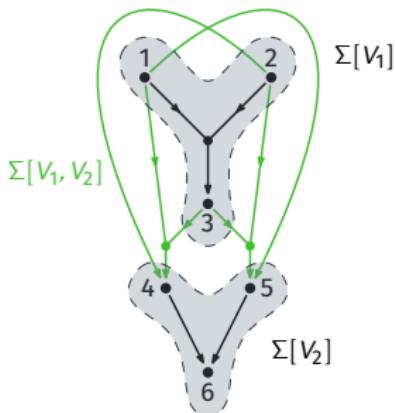
Definition ▷ Acyclic split

Σ an implicational base over V :

- *Acyclic* split of Σ : split (V_1, V_2) s.t. $A \subseteq V_1$ for each $A \rightarrow B \in \Sigma[V_1, V_2]$.

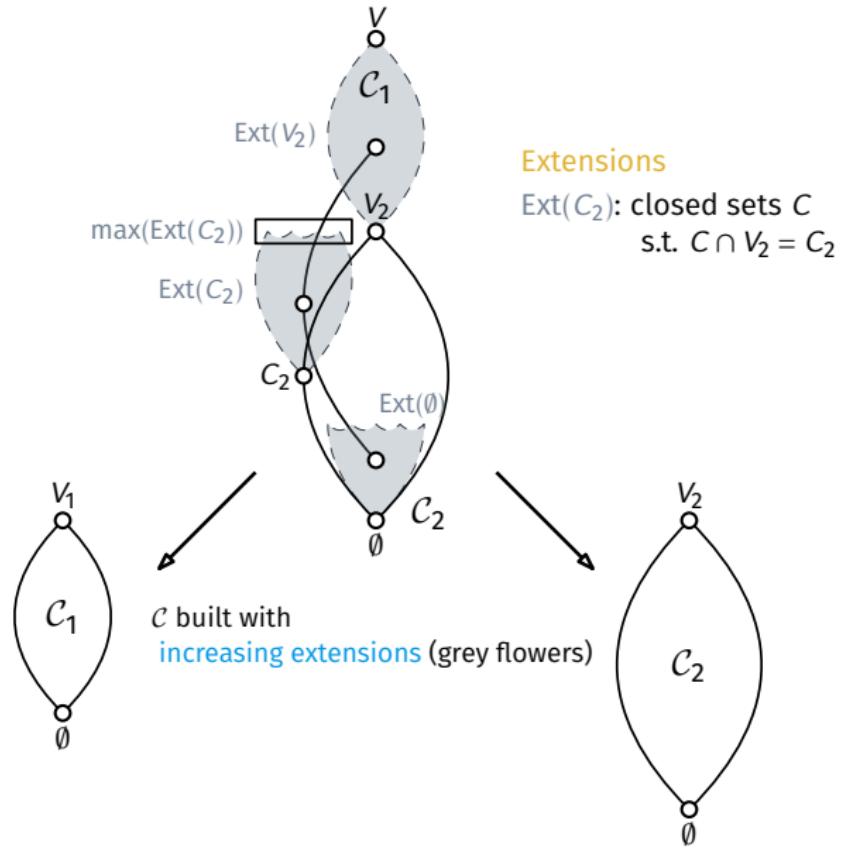


Cyclic split

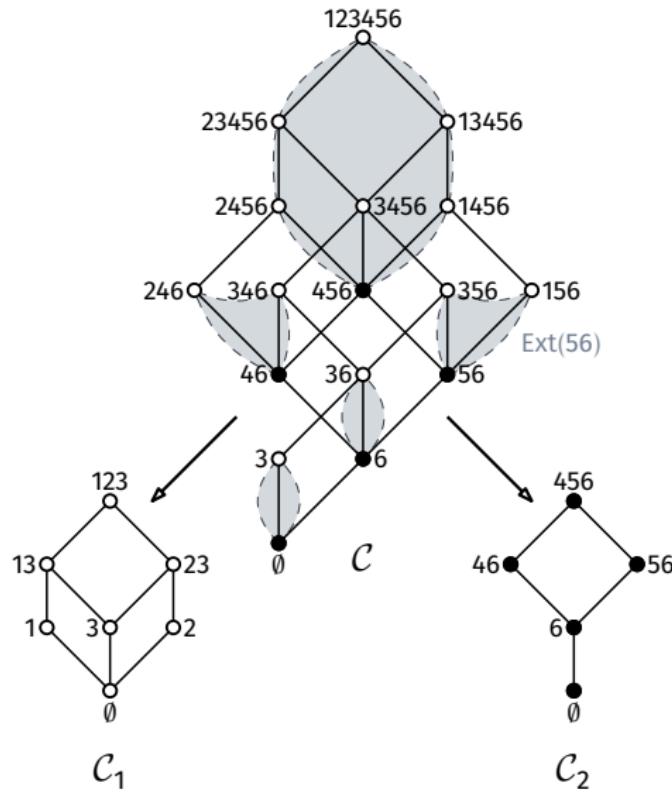


Acyclic split

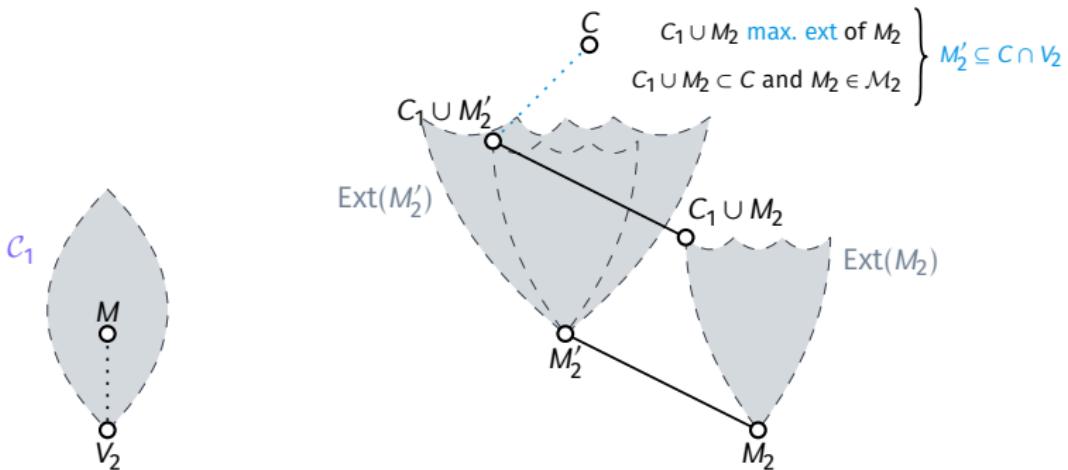
Translation ▷ Constructing \mathcal{C} with acyclic split



Translation ▷ Running example



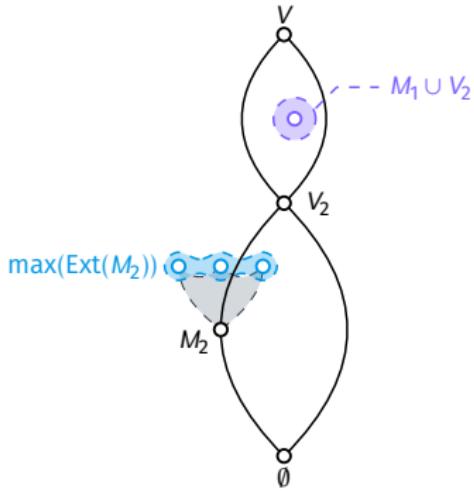
Translation ▷ Constructing \mathcal{M}



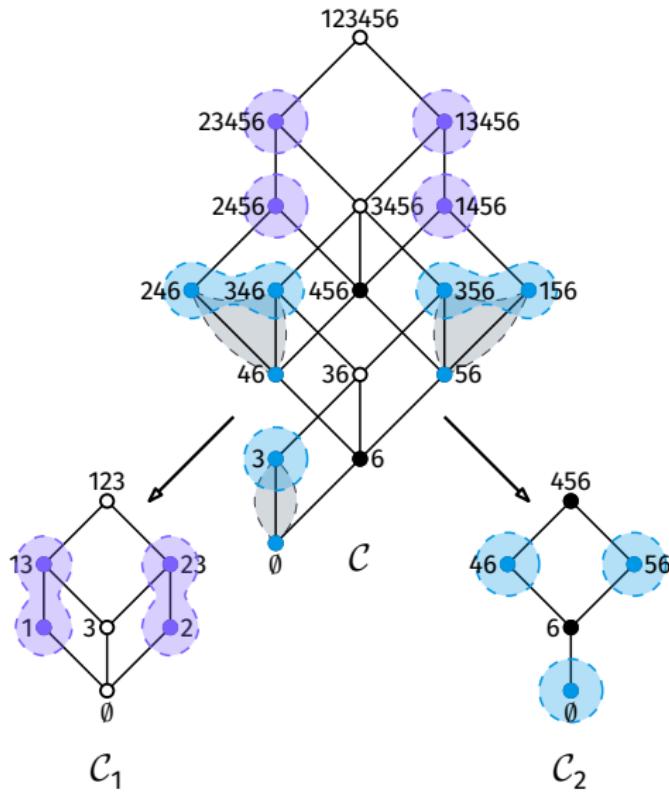
- Case 1: $V_2 \subseteq M$:
 - $C \in \text{Ext}(V_2)$ iff $C = C_1 \cup V_2$ ($C_1 \in \mathcal{C}_1$).
 - $M \in \mathcal{M}$ iff $M = M_1 \cup V_2$ ($M_1 \in \mathcal{M}_1$).
- Case 2: $V_2 \not\subseteq M$:
 - $M \in \text{Ext}(M_2)$, $M_2 \in \mathcal{M}_2$ (increasing extensions).
 - $M \in \max(\text{Ext}(M_2))$ for some $M_2 \in \mathcal{M}_2$.
 - $M \in \mathcal{M}$ iff $M \in \max(\text{Ext}(M_2))$ ($M_2 \in \mathcal{M}_2$).

Let Σ be an implicational base over V with acyclic split (V_1, V_2) . Then $|\mathcal{M}| \geq |\mathcal{M}_1| + |\mathcal{M}_2|$ and:

$$\mathcal{M} = \{M_1 \cup V_2 \mid M_1 \in \mathcal{M}_1\} \cup \{C \in \max(\text{Ext}(M_2)) \mid M_2 \in \mathcal{M}_2\}$$



Translation ▷ Running example



Translation ▷ Algorithm for CCM

Algorithm $\text{FindMeet}(\Sigma, V)$

Find an **acyclic** split (V_1, V_2) of Σ

If there is none:

 Compute \mathcal{M} with another algorithm

Else:

$\mathcal{M}_1 = \text{FindMeet}(\Sigma[V_1], V_1)$

$\mathcal{M}_2 = \text{FindMeet}(\Sigma[V_2], V_2)$

$\mathcal{M} = \{\mathcal{M}_1 \cup V_2 \mid \mathcal{M}_1 \in \mathcal{M}_1\}$

 For each $M_2 \in \mathcal{M}_2$:

$\mathcal{M} = \mathcal{M} \cup \max(\text{Ext}(M_2))$

Return \mathcal{M}

- **Beware:**

1. Size of $\mathcal{M}_1, \mathcal{M}_2$? ✓
2. Complexity of **ComputeMeet**?
3. Complexity of **finding extensions**.

Problem ▷ Computing Maximal Extension (MaxExt)

- *Input:* implicational base Σ with acyclic split (V_1, V_2) , \mathcal{M}_1 (resp. \mathcal{M}_2) the meet-irreducible elements associated to $\Sigma[V_1]$ (resp. $\Sigma[V_2]$), a closed set C_2 of $\Sigma[V_2]$.
 - *Output:* $\max(\text{Ext}(C_2))$.
-
- $\max(\text{Ext}(C_2))$ has a *dual* antichain in \mathcal{C}_1 coded by $\Sigma[V_1, V_2]$.
 - MAXEXT is then *equivalent* to dualization with \mathcal{M} *and* Σ .
 - If $\Sigma[V_1] = \emptyset$, MAXEXT is equivalent to MISENUM.

Corollary ▷ Applications

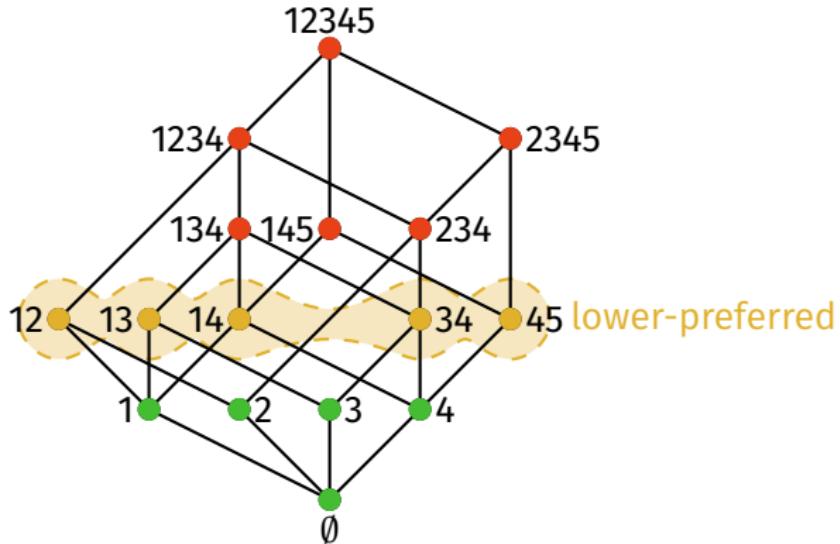
Let Σ be an implicational base over V . Assume there exists a full partition V_1, \dots, V_k of V such that for every implication $A \rightarrow b \in \Sigma$, $A \subseteq V_i$ and $b \in V_j$ for some $1 \leq i < j \leq k$. Then CCM can be solved in *output-quasipolynomial time*.

- Particular case of *acyclic convex geometry* [Adaricheva, 2017], [Hammer, Kogan, 1995].
- Generalizes *ranked convex geometry* [Defrain, Nourine, V., 2021], where CCM is equivalent to MISENUM.
- Also works for “*simple closure systems*” (diamonds, pentagons, etc).

- Problem:
 - CCM: enumerating meet-irreducible elements from implications.
 - *Unknown complexity, harder than MISENUM.*
- Results:
 - (Acyclic) split operation.
 - Hierarchical decomposition of Σ , recursive construction of \mathcal{M} .
 - New tractable cases (*output-quasipolynomial time*) in acyclic convex geometries.
- Further research:
 - *Recognition* of an acyclic split from \mathcal{M} ?
 - Generalization to “*simple*” *non-acyclic splits*?
 - Complexity of CCM in (acyclic) convex geometries?

Second problem ► Forbidden pairs in closure systems

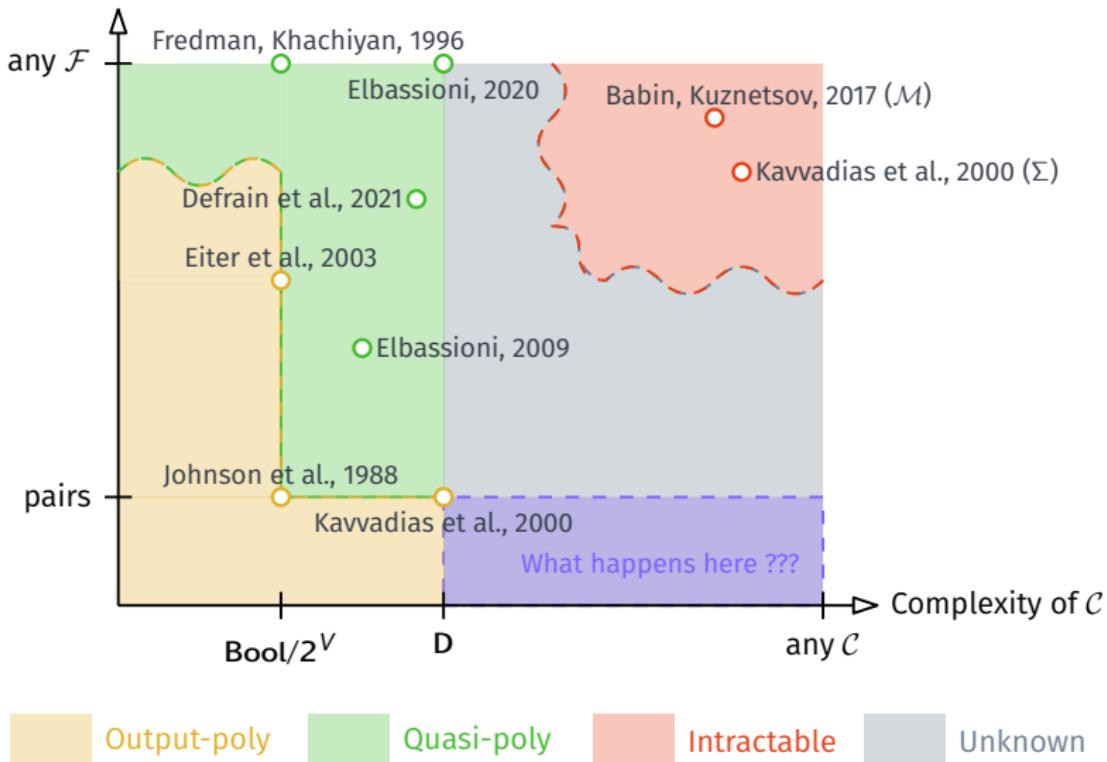
Forbidden pairs ▷ Dualization and forbidden sets



- Closure system \mathcal{C} (given by Σ or \mathcal{M}), *forbidden sets* $\mathcal{F} = \{134, 145, 24\}$.
- 1 is *lower-admissible* : does not contain a set in \mathcal{F} .
- 12 is *lower-preferred* : *inclusion-wise max.* lower-admissible.

Forbidden pairs \triangleright The hardness of dualization DUAL(α)

Complexity of \mathcal{F}



Definition ▷ lower-preferred closed sets

\mathcal{C} closure system over V , family \mathcal{F} over V of *forbidden pairs* for \mathcal{C} :

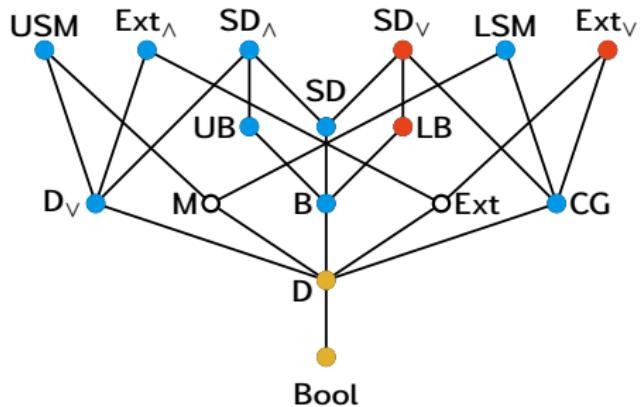
- *Lower-admissible* if $F \not\subseteq C$ for each $F \in \mathcal{F}$.
- *Lower-preferred* if inclusion-wise max. lower-admissible.

Problem ▷ Enum. Lower-Pref. with forb. Pairs (ELP-P(α))

- *Input:* a representation α for a closure system \mathcal{C} , a family \mathcal{F} of *forbidden pairs* (both over V).
- *Output:* *lower-preferred closed sets* of \mathcal{C} w.r.t. \mathcal{F} .

- Models inconsistency:
 - *Poset + forbidden pairs*: representation for median semilattices [Barthélemy, Constantin, 1993].
 - *implications + forbidden pairs*: representation for modular semilattices [Hirai, Nakashima, 2018].

Forbidden pairs ▷ The complexity of ELP-P(α)



- Bool = Boolean
- D = Distributive
- SD = Semidistributive
- Ext = Extremal
- M = Modular
- SM = Semimodular
- B = Bounded
- CG = Convex Geometry

- Output-poly
- $\geq \text{DUAL}(\alpha)$ in CG, B or D_V
- ELP-P(Σ) hard

Corollary ▷ Nourine, V.

The problem ELP-P(α) is *intractable*. Moreover, ELP-P(Σ) is *intractable* in *lower-bounded* and *join-extremal* closure systems.

Definition ▷ Minimal Generator, Carathéodory number

\mathcal{C} (standard) closure system over V :

- $A \subseteq V$ *minimal generator* of $u \in V$: $u \in \phi(A)$ and $u \notin \phi(A')$, $\forall A' \subset A$.
- *Carathéodory number* $\text{cc}(\mathcal{C})$ of \mathcal{C} : *maximal size* of a minimal generator.

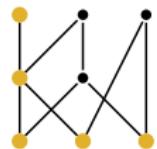
Theorem ▷ Nourine, V.

The problem ELP-P(α) can be solved in:

- *Output-polynomial time* if $\text{cc}(\mathcal{C}) \leq k$, for some constant $k \in \mathbb{N}$.
- *Output-quasipolynomial time* if $\text{cc}(\mathcal{C}) \leq \log(|V|)$.

Forbidden pairs ▷ Tractable cases

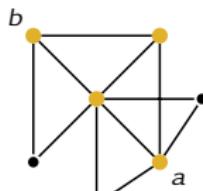
- Closure systems where $\text{cc}(\mathcal{C})$ is constant:



ideals of a poset
 $\text{cc}(\mathcal{C}) = 1$



convex subsets of a poset
 $\text{cc}(\mathcal{C}) \leq 2$



geod. conv. of a chordal graph
 $\text{cc}(\mathcal{C}) \leq 2$



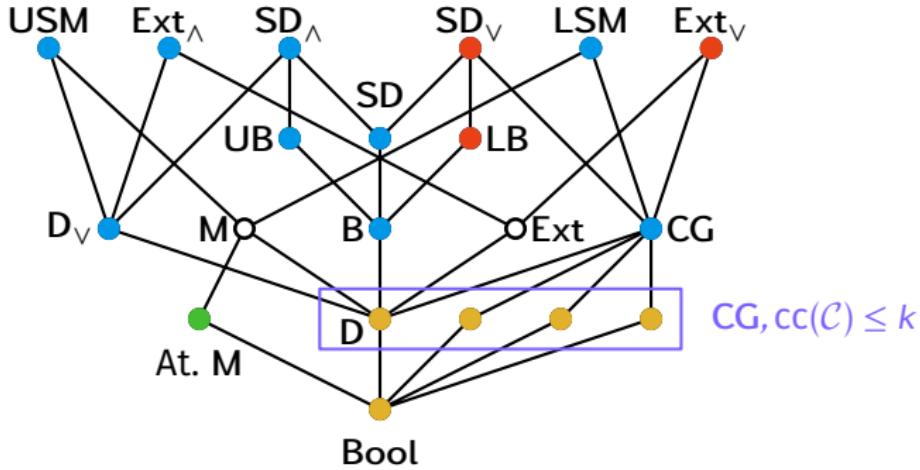
convex hull in \mathbb{R}^k
 $\text{cc}(\mathcal{C}) = k + 1$

- Biatomicity [Bennett, 1987] + Independence criterion [Grätzer, 2011].

Corollary ▶ Nourine, V.

The problem ELP-P(α) can be solved in *output-quasipolynomial time* in atomistic modular closure systems.

Forbidden pairs $\triangleright \text{ELP-P}(\alpha)$: the big picture



- Output-poly
- Quasi-poly
- $\geq \text{DUAL}(\alpha)$
- $\text{ELP-P}(\Sigma)$ hard in D_V , B , or CG

- Further research:
 - Complexity of $\text{ELP-P}(\alpha)$ in *modular* and *extremal* closure systems?
 - Characterize the lattices where $\text{ELP-P}(\alpha) \equiv$ enumerate max. independent sets of a graph?

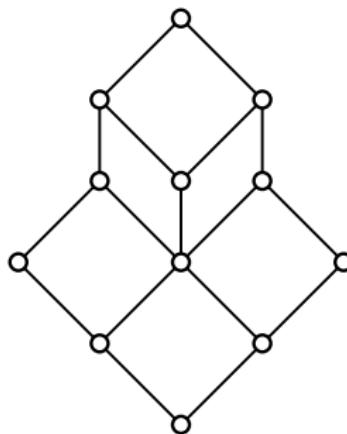
Conclusion ▷ Summary and perspectives

- Context:
 - Initial motivation from knowledge spaces (ProFan project).
 - Theoretical study of closure systems and their representations.
- First problem – translating between the representations:
 - *Unknown complexity, harder than MISENUM.*
 - *New tractable classes* based on *hierarchical decompositions* of implications.
 - *(Not in this talk)* previous work on *ranked convex geometries*.
- Second problem – closure systems with forbidden sets:
 - Enumerating admissible and preferred closed sets.
 - *Hardness results for ELP-P(α)* using DUAL(α), *tractable cases* based on the Carathéodory number.
 - *(Not in this talk)* results for *forbidden supersets*.
- Open questions:
 - What is the complexity of CCM in acyclic convex geometries?
 - Characterize the lattices where $\text{ELP-P}(\alpha) \equiv \text{max. independent sets of a graph}$?

Conclusion ▷ Productions

- Translation:
 - *The enumeration of meet-irreducible elements based on hierarchical decompositions of implicational bases.* With Lhouari Nourine.
Submitted to *Theoretical Computer Science* and communicated at *WEPA 2020, FCA4AI 2020, ICTCS 2020*.
 - *Translating between the representations of a ranked convex geometry.* With Oscar Defrain and Lhouari Nourine.
Published in *Discrete Mathematics* (2021) and communicated at *WEPA 2019*.
- Forbidden sets:
 - *Enumerating maximal consistent closed sets in closure systems.* With Lhouari Nourine.
Published in *Proceedings of ICFCA 2021* and communicated at *ICFCA 2021*.
- Other:
 - *Towards declarative comparabilities: application to functional dependencies.* With Lhouari Nourine and Jean-Marc Petit.
Under review in *Journal of Computer and System Sciences* and communicated at *BDA 2021*.

Thank you for your attention!



Conclusion ▷ References

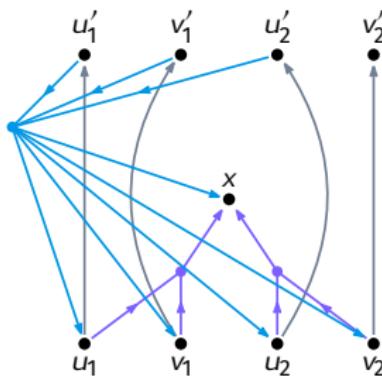
- ▶ **K. Adaricheva**
Optimum Basis of Finite Convex Geometry.
Discrete Applied Mathematics, 230 :11-20, 2017.
- ▶ **M. Babin, S. Kuznetsov**
Dualization in lattices given by ordered sets of irreducibles.
Theoretical Computer Science, 658 :316-326, 2017.
- ▶ **J-P. Barthélemy, J. Constantin**
Median graphs, parallelism and posets.
Discrete Mathematics, 111 :49-63, 1993.
- ▶ **L. Beaudou, A. Mary, and L. Nourine.**
Algorithms for k -meet-semidistributive lattices.
Theoretical Computer Science, 658 :391-398, 2017.
- ▶ **M.K. Bennett**
Biatomic lattices.
Algebra Universalis, 24 :60-73, 1987.
- ▶ **K. Bertet, C. Demko, and 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.
- ▶ **O. Defrain, L. Nourine, T. Uno.**
On the dualization in distributive lattices and related problems.
Discrete Applied Mathematics, 300 :85–96, 2021.

- ▶ O. Defrain, L. Nourine, S. Vilmin.
Translating between the representations of a ranked convex geometry.
Discrete Mathematics, 344 :112399, 2021.
- ▶ J.P. Doignon, J.C. Falmagne.
Spaces for the assessment of knowledge.
International journal of man-machine studies, 23(2) :175-196, 1985.
- ▶ T. Eiter, G. Gottlob, K. Makino.
New results on monotone dualization and generating hypergraph transversals.
SIAM Journal on Computing, 32 :514-537, 2003.
- ▶ K. Elbassioni.
Algorithms for dualization over products of partially ordered sets.
SIAM Journal on Discrete Mathematics, 23 :487-510, 2009.
- ▶ K. Elbassioni.
On dualization over distributive lattices.
arXiv preprint, arXiv:2006.15337, 1996.
- ▶ M. Fredman, L. Khachiyan.
On the complexity of dualization of monotone disjunctive normal forms.
Journal of Algorithms, 21 :618-628, 1996.
- ▶ G. Grätzer
Lattice theory: foundation.
Springer Science & Business Media, 2011.

- ▶ P. Hammer, A. Kogan
Quasi-Acyclic Propositional Horn Knowledge Bases: Optimal Compression.
IEEE Transactions on knowledge and data engineering, 7 :751–762, 1995.
- ▶ H. Hiroshi and T. Oki
A compact representation for minimizers of k -submodular functions.
Journal of Combinatorial Optimization, 36 :709–741, 2018.
- ▶ R. Khardon.
Translating between Horn Representations and their Characteristic Models.
Journal of Artificial Intelligence Research, 3 :349–372, 1995.
- ▶ D. Kavvadias, M. Sideri, and E. Stavropoulos
Generating all maximal models of a Boolean expression.
Information Processing Letters, 74 :157–162, 2000.
- ▶ D. Johnson, M. Yannakakis, and C. Papadimitriou.
On Generating All Maximal Independent Sets.
Information Processing Letters, 27 :119–123, 1988.
- ▶ H. Mannila, K.-J. Räihä.
The design of relational databases.
Addison-Wesley Longman Publishing Co., Inc., 1992.
- ▶ M. Wild.
The Joy of Implications, Aka Pure Horn Formulas: Mainly a Survey.
Theoretical Computer Science, 658 :264–292, 2017.
- ▶ M. Wild.
Computations with Finite Closure Systems and Implications.
Springer LNCS, 959, 1995.

Appendix ▷ Beyond acyclic splits

- universe $V = V_1 \cup V_2$ with:
 - $V_1 = \{u_1, \dots, u_n, v_1, \dots, v_n, x\}$, $V_2 = \{u'_1, \dots, u'_n, v'_1, \dots, v'_n\}$, $n \in \mathbb{N}$
- Σ over V with split (V_1, V_2) :
 - $\Sigma[V_1] = \{u_i v_i \rightarrow x \mid 1 \leq i \leq n\}$, $\Sigma[V_2] = \emptyset$
 - $\Sigma[V_1, V_2] = \{u_i \rightarrow u'_i \mid 1 \leq i \leq n\} \cup \{v_i \rightarrow v'_i \mid 1 \leq i \leq n\} \cup \{A \rightarrow V_1 \mid A \subseteq V_2, |A| = 3\}$



Appendix ▷ Beyond acyclic splits

