THE D-BASE OF FINITE CLOSURE SYSTEMS

Simon Vilmin (LIS, Aix-Marseille Université)

Journées smartFCA - 21/06/24

Joint work with:

Kira Adaricheva (Math dpt., Hofstra University)

Lhouari Nourine (LIMOS, Université Clermont Auvergne)

(arXiv link)

Who am I

2013 ISIMA : IT engineering school 2018

Université Clermont Auvergne (UCA)

10/18 PhD : lattices, enumeration 12/21

LIMOS, UCA, with Lhouari Nourine

02/22 Post-doc : databases, functional dep. 01/23

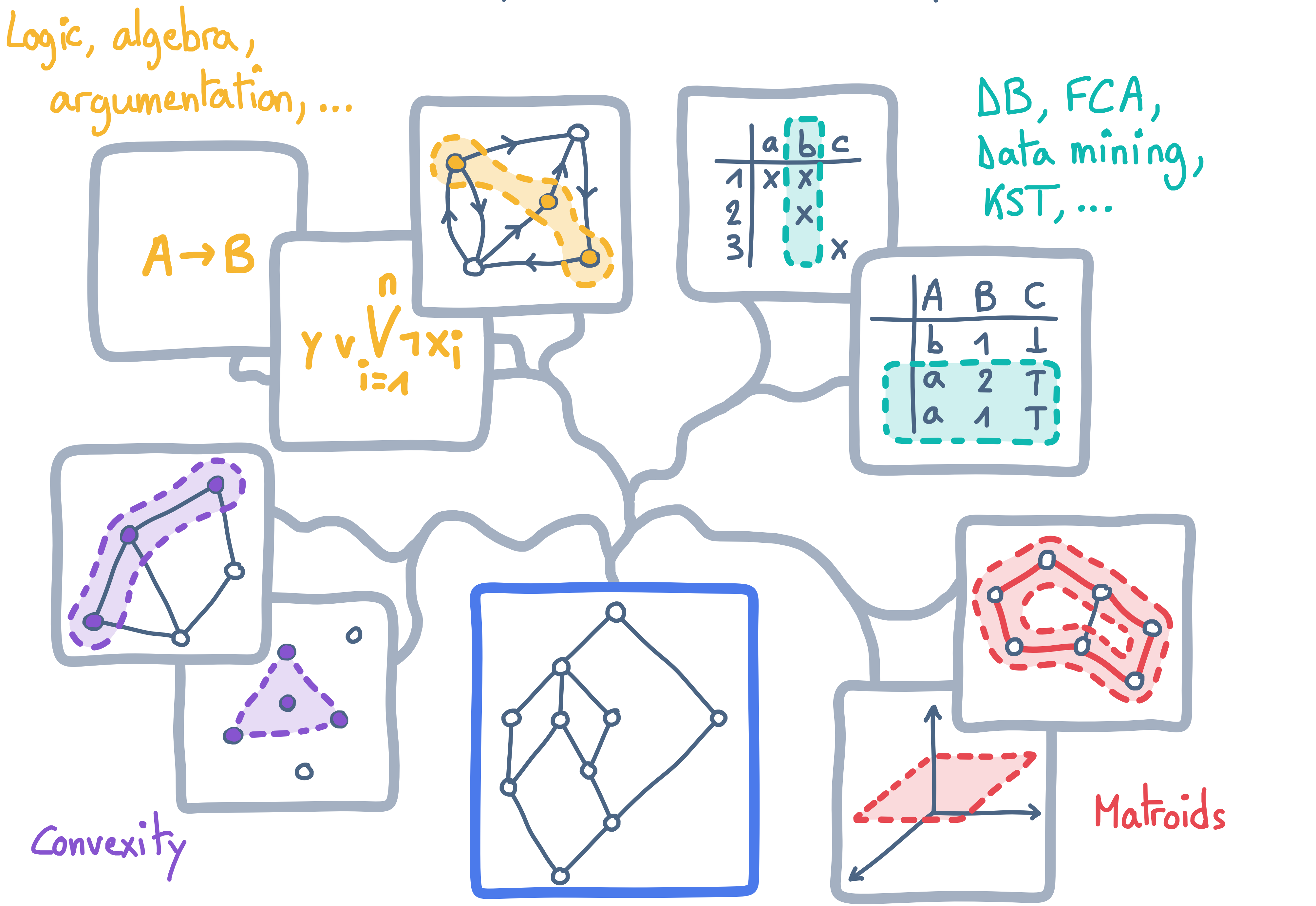
LIRIS, INSA Lyon, with Jean-Marc Petit

03/23 Post-doc : median lattices 08/23

LORIA, with Amedeo, Miguel

09/23 MCF LIS

Research : representations of closure systems and lattices



Outline

PART I: what is the D-base

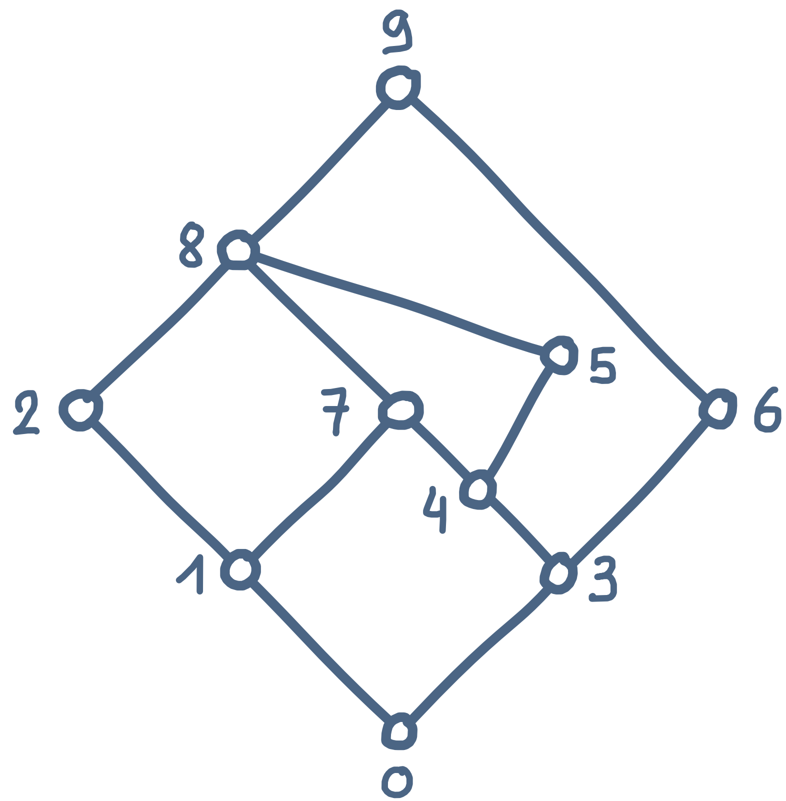
- from lattices ...
- ... to closure systems (and minimal generators)

PART II: computing the D-base

- from implications
- from meet-irreducible elements

PART I: What is the D-base

starting point: lattices and join covers



IDEA: describe the (structure of the) lattice by means of join covers of join-irreducible elements

A long-standing question

Dilworth, 40 : unique irreducible decompositions in locally distributive lattices ($\simeq$ convex geometries)

Finkbeiner, 51 : dependence relation (= implications)

Jónsson, Nation, 77 : join-refinement, minimal join covers (D-base) for free lattices

Gaskill, Rival, 78 : use of "minimal pairs" in modular lattices (= D-base)

A long-standing question

Day, 79 : study of (lower, upper) bounded lattices with relations on join-covers (D-relation, D- base)

Faigle, 86 : minimal pairs in geometric lattices

Nation, 90 : OD - graph (D-base) — in this talk

Freese et al., 95 : book on free lattices

in this talk

Bertet, Monjardet, 10 : survey on minimal generators and canonical direct base ($\supseteq$ D - base)

Leading to the Δ-base

in this talk

Adaricheva et al., 13 : introduction of the Δ-base

Rodriguez et al., 15, 17 : computing Δ-base with simplification logic (from implications)

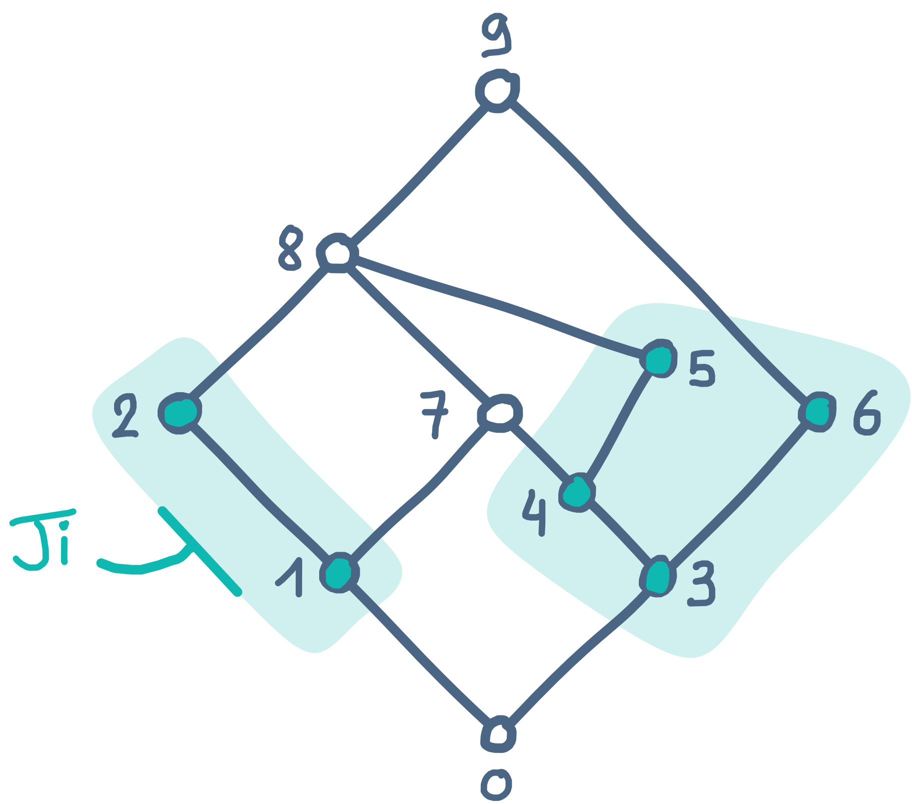
Adaricheva, Nation, 17 : computing Δ-base with hypergraph dualization (from context)

In this talk

Adaricheva, Nourine, Vilmin, 24+ : output-sensitive study of Δ-base computation, with implications or meet-irreducibles (= context)

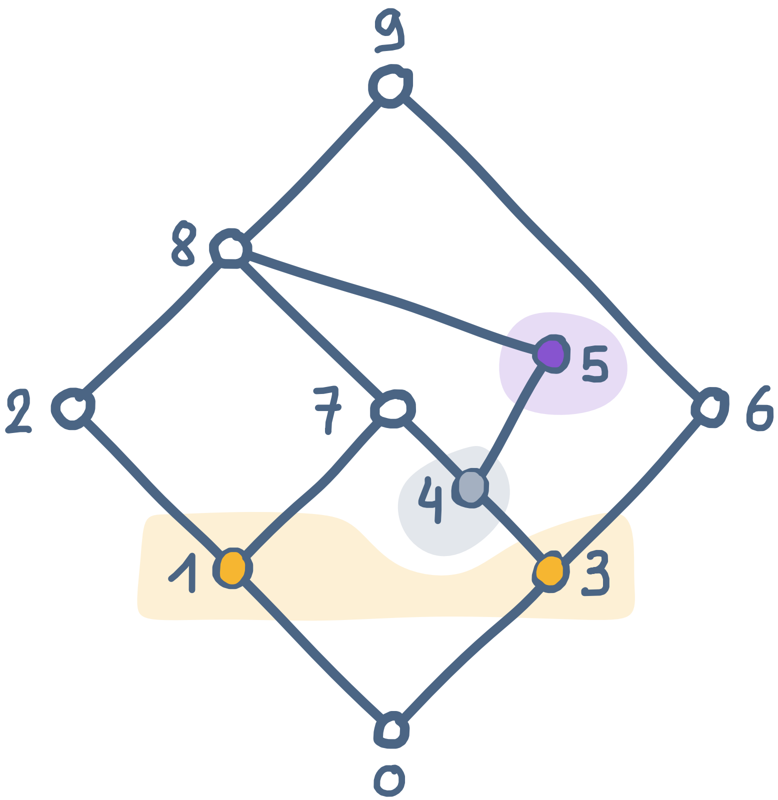
DEF (join-irr): in a lattice $\mathcal{L}$, $j \in \mathcal{L}$ is join-irreducible if $j \neq$ bot and $j = a \vee b$ entails $j = a$ or $j = b$, for $a, b \in \mathcal{L}$.

$$\mathcal{J}i = \{ j \in \mathcal{L} : j \text{ join-irreducible} \}$$



DEF (join-cover): given $j \in J_i$, a join-cover of j is a subset A of J_i such that $\boxed{j \leqslant \bigvee A}$

implication $A \to j$



Two "types" of join covers

- trivial, based on $(J_i, \leqslant)$

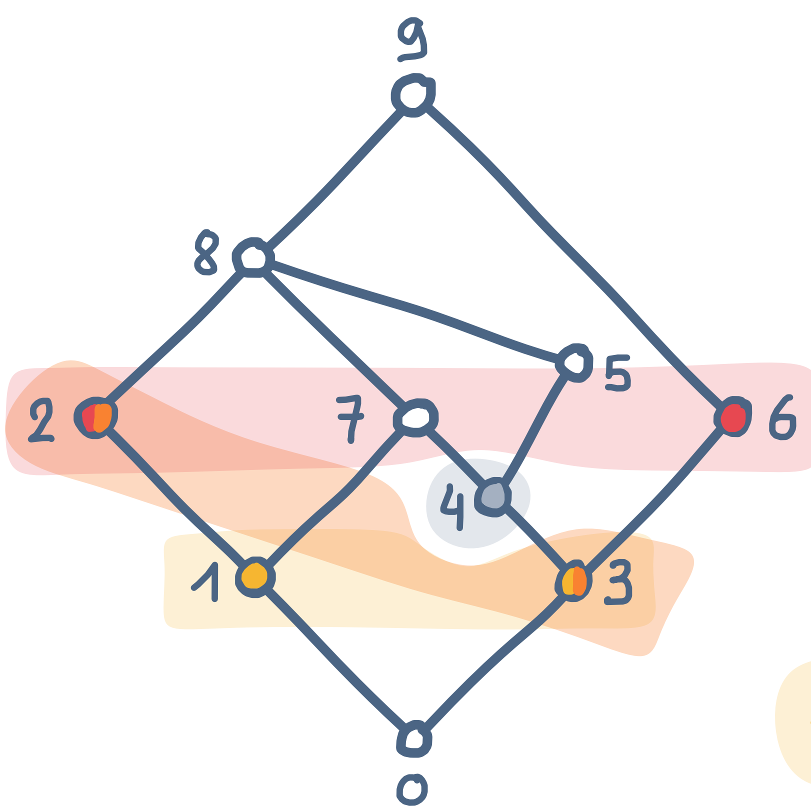
  $4 \leqslant 5$

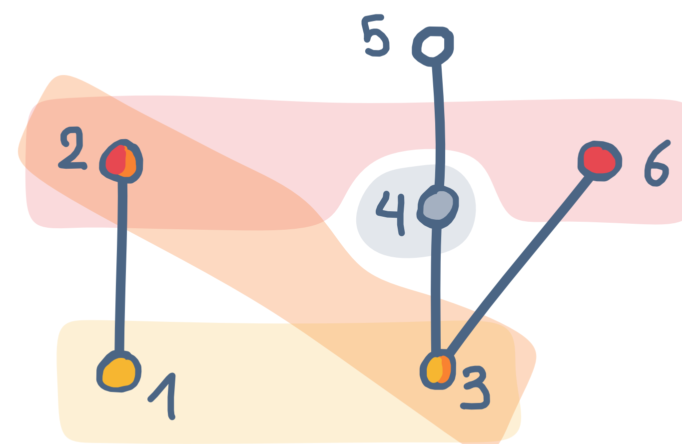
- non-trivial, $j \not\leqslant a$ for $a \in A$

  $4 \leqslant 1 \vee 3$

Minimal non-trivial join-cover : intuition

IDEA : among non-trivial join-covers of j, keep the lowest ones w.r.t. $\ll$ (or $(J_i, \leq)$ equivalently)





$(J_i, \leq)$

13 lower than 23 lower than 26 among covers of 4

Minimal non-trivial join-cover: formalization

DEF (refinement): given $A, B \subseteq J_i$, A **refines** B if for each $a \in A$, there is some $b \in B$ s.t. $a \leq b$

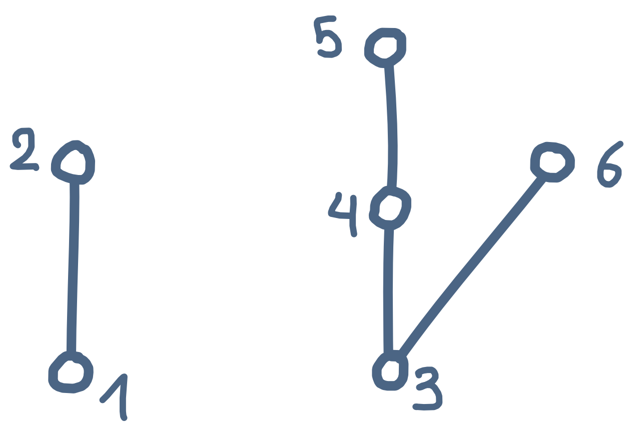
DEF (D-generator): a **minimal non-trivial join-cover** of j, or **D-generator of j**, is a ntjc of j that **cannot be refined** to another ntjc of j.

RMK:

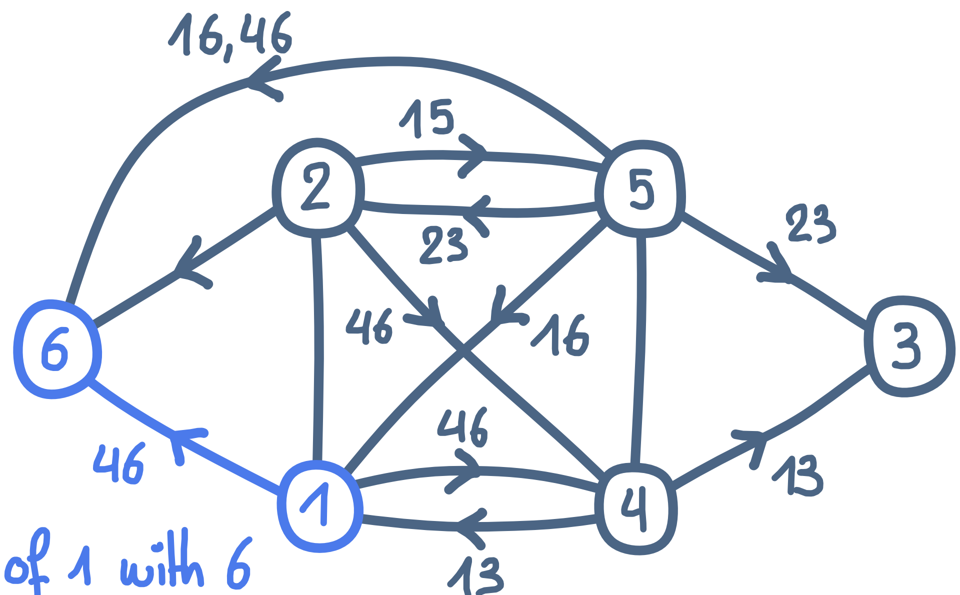
- D-generators: minimal join-cov w.r.t. refinement
- $A_1 \subseteq A_2$ implies that A_1 refines A_2 !

OΔ-graph Nation, 90

IDEA: O order of join irreducible $(J_i, \leq)$ + Δ directed labeled (multi-)graph based on Δ-generators



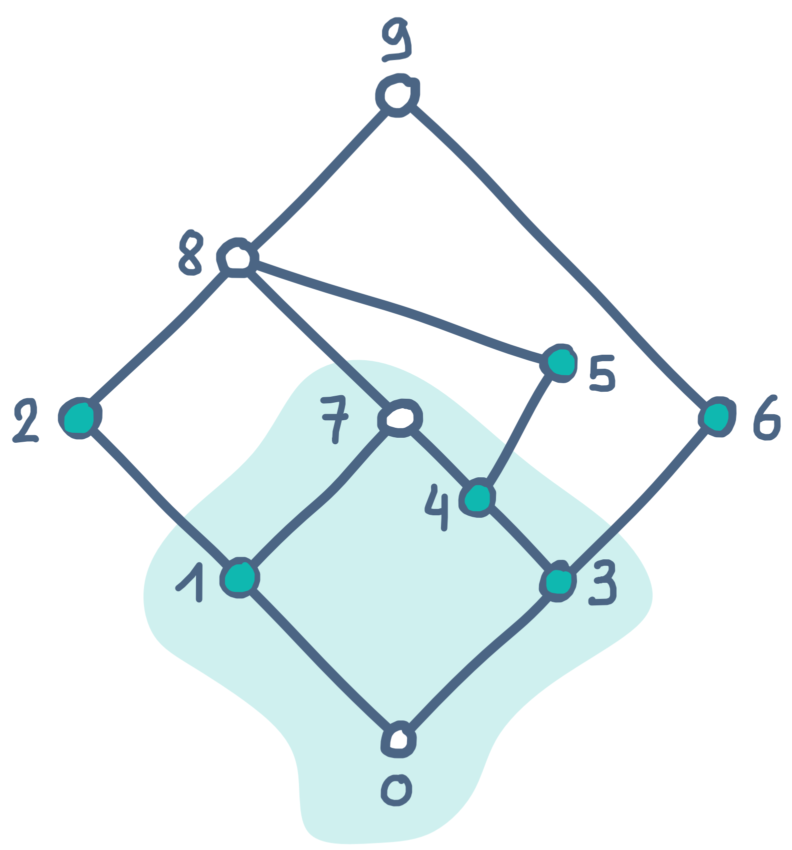
$(J_i, \leq)$

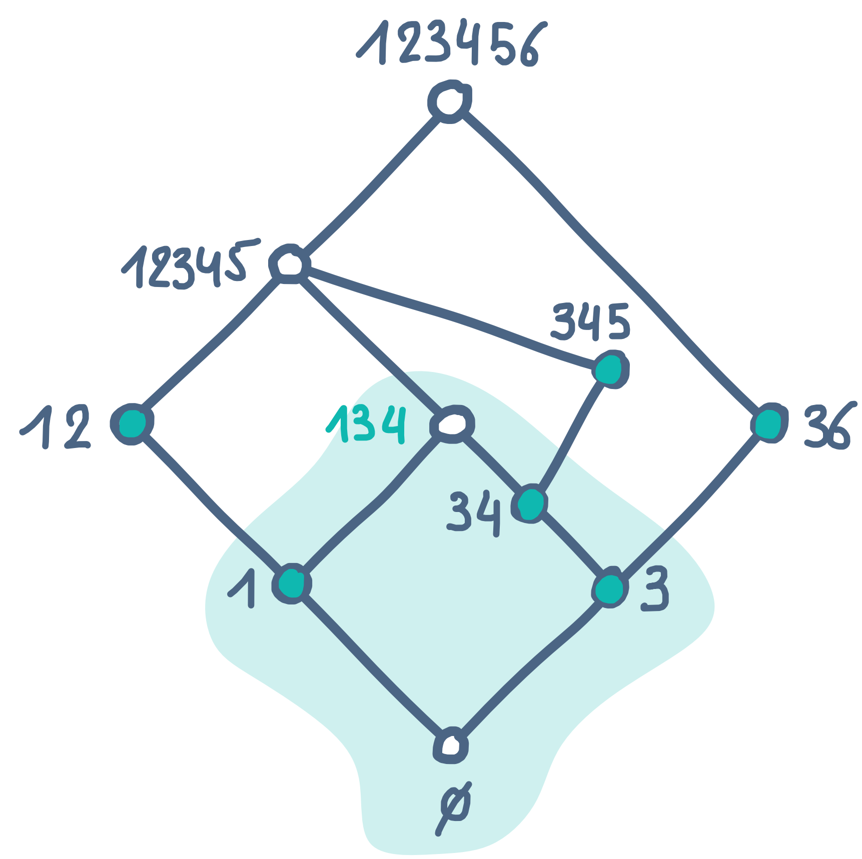


46 Δ-gen of 1 with 6

allows to rebuild the lattice as a closure system

From lattices to closure systems





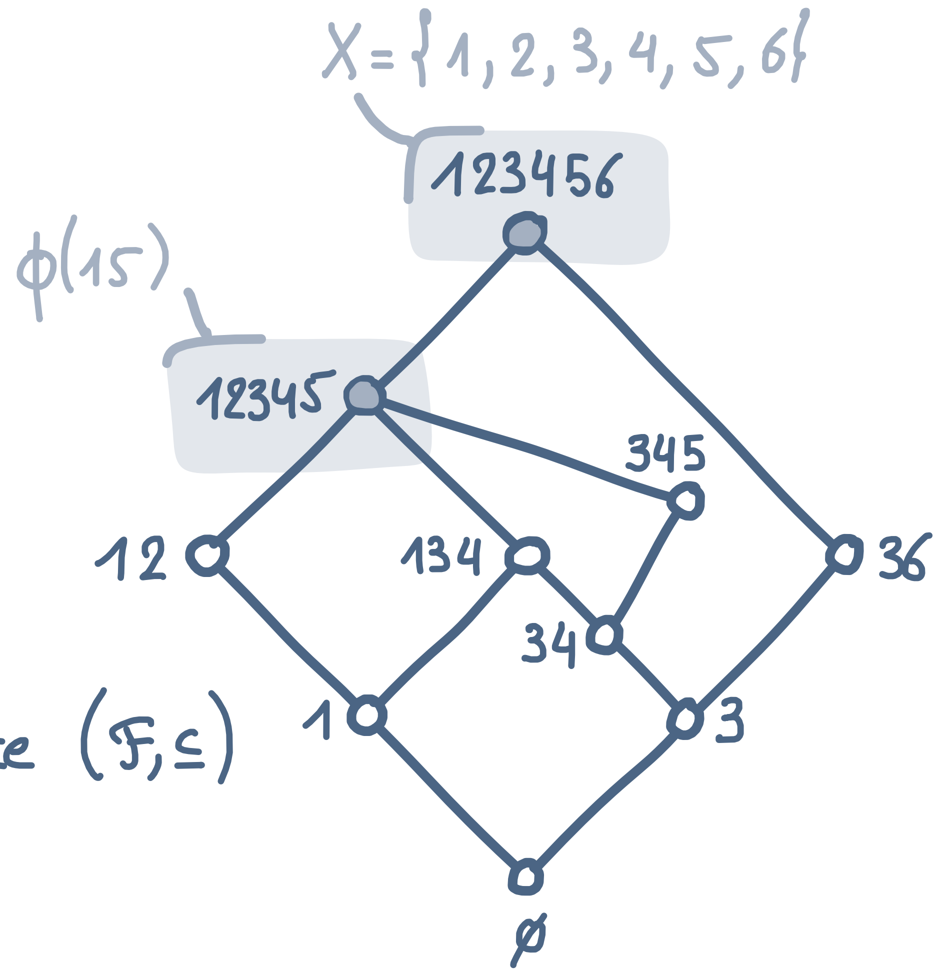
$4 \leq 5,\ 4 \leq 1 \vee 3 \longrightarrow 5 \to 4,\ 13 \to 4$

DEF (closure system): set system $\mathcal{F}$ over groundset X s.t. $X \in \mathcal{F}$ and $F_1 \cap F_2 \in \mathcal{F}$ for $F_1, F_2 \in \mathcal{F}$.

Sets in $\mathcal{F}$: closed sets

$\mathcal{F}$ induces a closure operator:

$$\phi(A) = \bigcap \{F \in \mathcal{F} : A \subseteq F\}$$

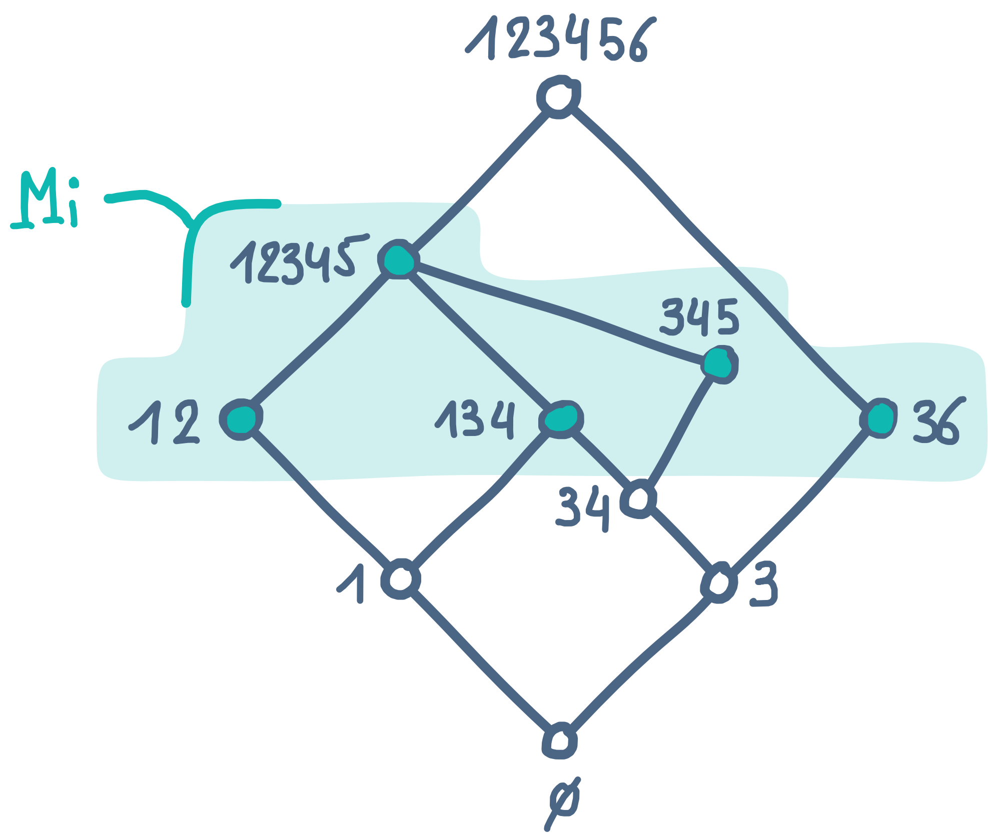


closure lattice $(\mathcal{F}, \subseteq)$

Meet-irreducible closed sets

DEF (meet-irreducible): a closed set $M \in \mathcal{F}$ distinct from X is (meet-)irreducible if for all $F_1, F_2 \in \mathcal{F}$ $M = F_1 \cap F_2$ entails $M = F_1$ or $M = F_2$

$$\underline{Mi} = \{ M \in \mathcal{F} : M \text{ is irreducible} \}$$



Implications

(or $A \to B$, $B \subseteq X$) "If I have A, I have b"

DEF (implications):

- **implication**: statement $\boxed{A \to b}$ with $A \subseteq X$, $b \in X$
- **implicational base (IB)**: set Σ of implications on X

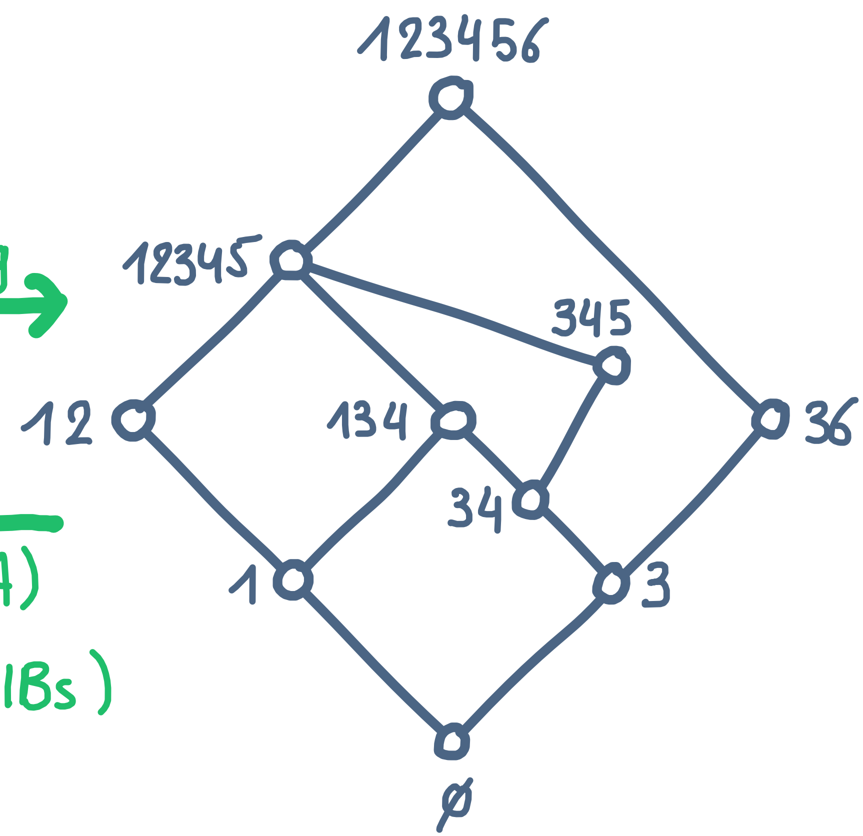
Binary implications

$$\Sigma = \begin{bmatrix} 2 \to 1, & 4 \to 3, \\ 5 \to 4, & 6 \to 3, \\ 13 \to 4, & 23 \to 5, \\ 15 \to 2, & 46 \to 2, \\ 16 \to 5 & \end{bmatrix}$$

forward chaining $\longrightarrow$

$\longleftarrow$ $b \in \phi(A)$

(many equiv. IBs)



Minimal generators, canonical direct base Bertet, Monjardet, 10

DEF (minimal generator): $A \subseteq X$ is a minimal generator of x if it is an $\subseteq$-min subset of X satisfying $x \in \phi(A)$

The canonical direct base of a closure system is

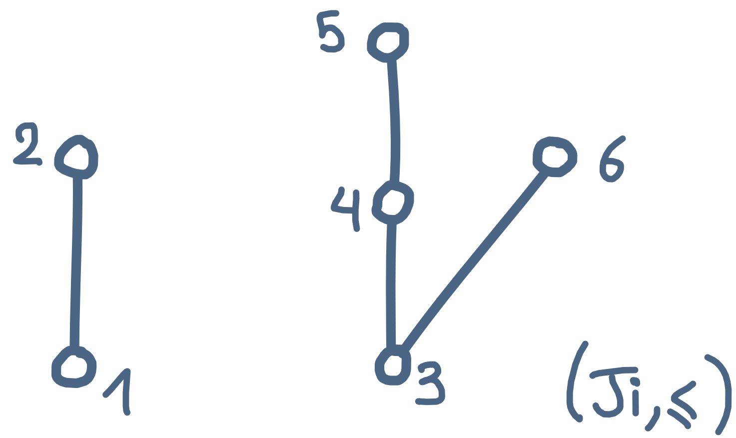
$$\Sigma_{cd} = \{ A \to x : A \text{ minimal generator of } x,\ x \notin A \}$$

Our running example:

$$\Sigma_{cd} = \begin{bmatrix} 2 \to 1,\ 46 \to 1,\ 56 \to 1, \\ 46 \to 2,\ 15 \to 2,\ 16 \to 2, \\ 5 \to 3,\ 4 \to 3,\ 6 \to 3, \\ 13 \to 4,\ 23 \to 4,\ 16 \to 4,\ 26 \to 4 \\ 14 \to 5,\ 16 \to 5,\ 24 \to 5,\ 26 \to 5 \end{bmatrix}$$

What about the D-base then ?

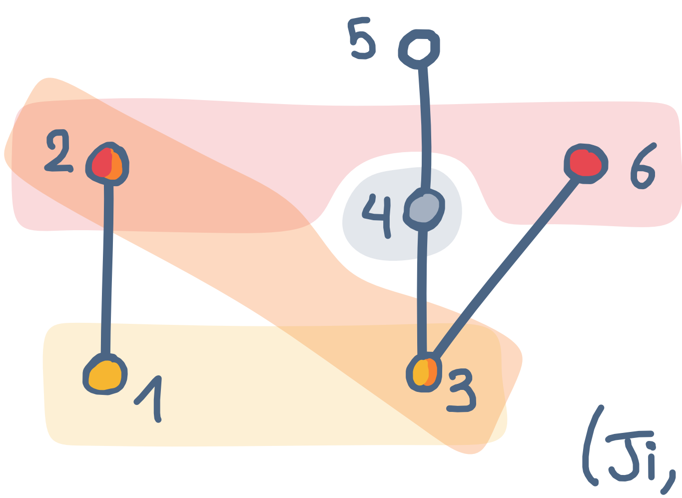
1st ingredient : order to binary implications



$$\Sigma_b = \left[2 \to 1,\ 5 \to 4,\ 5 \to 3,\ 4 \to 3,\ 6 \to 3 \right]$$

with closure operator ϕ_b

2nd ingredient : translate D-generators



non-singleton minimal generators of x minimal w.r.t. ϕ_b !

$13 \subseteq 123 \subseteq 1236$

$\phi_b(13) \subseteq \phi_b(23) \subseteq \phi_b(26)$

At last, the D-grail Adaricheva et al., 13

DEF (D-generator): $A \subseteq X$ is a D-generator of x if $x \notin \phi_b(A)$ and A is a minimal generator of x being ϕ_b-minimal among min. gen. of x

DEF (D-base): the D-base of a closure system is:

$$\Sigma_D = \Sigma_b \cup \{A \to x : A \text{ } D\text{-gen of } x\}$$

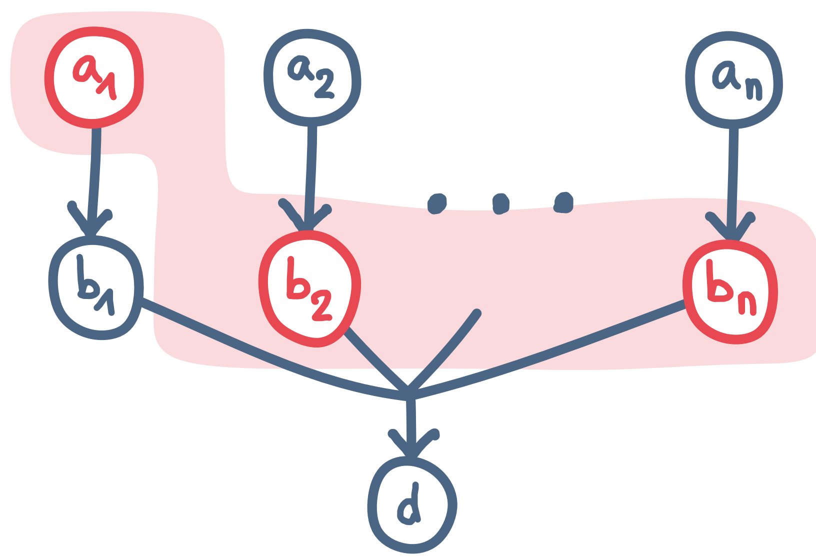
$$\Sigma_D = \begin{bmatrix} 2 \to 1, \ 5 \to 4, \ 5 \to 3, \\ 4 \to 3, \ 6 \to 3 \end{bmatrix} \cup \begin{bmatrix} 46 \to 1, \\ 46 \to 2, \ 15 \to 2, \ 16 \to 2, \\ 13 \to 4, \\ 14 \to 5, \ 16 \to 5 \end{bmatrix}$$

Bonus: gap between Σ_{cd} and Σ_Δ

RMK: we have $\Sigma_\Delta \subseteq \Sigma_{cd}$, but what is the gap?

$X = \{a_1, \dots, a_n, b_1, \dots, b_n, d\}$

$\Sigma = \{a_i \to b_i : 1 \leq i \leq n\} \cup \{b_1 \dots b_n \to d\}$

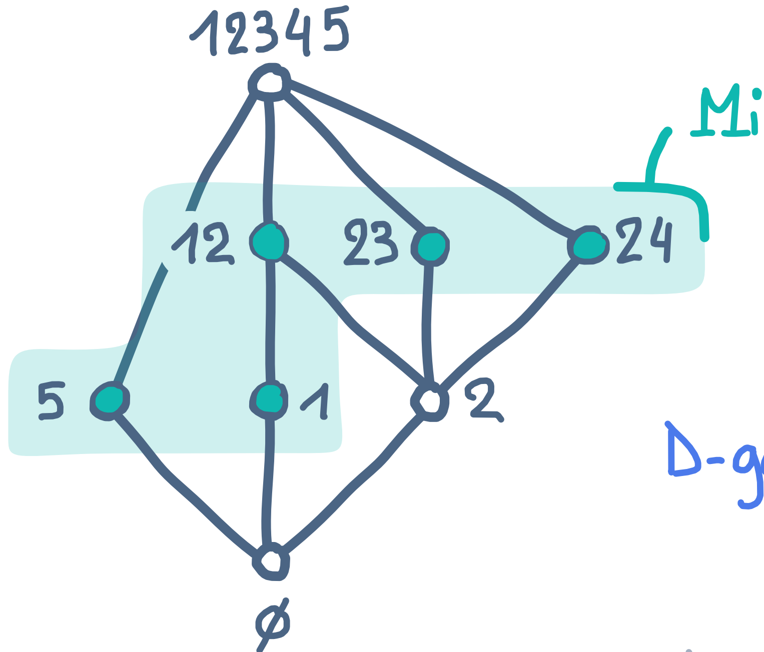


$\Sigma_\Delta = \Sigma \qquad \Sigma_{cd} = \Sigma_\Delta \cup \{A \to d : A \in \prod_{i=1}^{n} \{a_i, b_i\}\}$

exponential gap w.r.t. Σ_Δ!

PART II: Computing the Δ-base

Brand new toy example



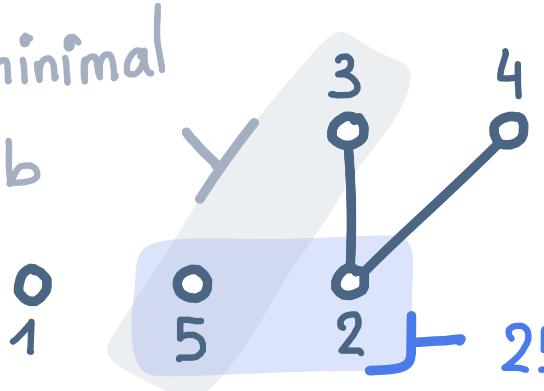
Binary implications Σ_b

$$\Sigma = \begin{bmatrix} 4 \to 2, & 25 \to 3, \\ 3 \to 2, & 34 \to 1, \\ 25 \to 4, & 14 \to 3, \\ 15 \to 4, & 13 \to 5, \\ 35 \to 4 & \end{bmatrix}$$

D-generator of 4

minimal generators of 4
(no subset of 35 gives 4)

35 not minimal w.r.t. Σ_b



Σ_b as a poset

25 minimal w.r.t. $\Sigma_b \Rightarrow$ D-generator of 4

PROB : given irreducibles M_i over X, find the D-base Σ_D

PROB : given an implicational base Σ over X, find the D-base Σ_D

RMK : enumeration tasks, listing implications without repetitions

Further motivations

Theoretical / algorithmic properties :

- convey structural information of closure systems
- ordered direct (fast forward chaining)
- much smaller than the set of all minimal generators

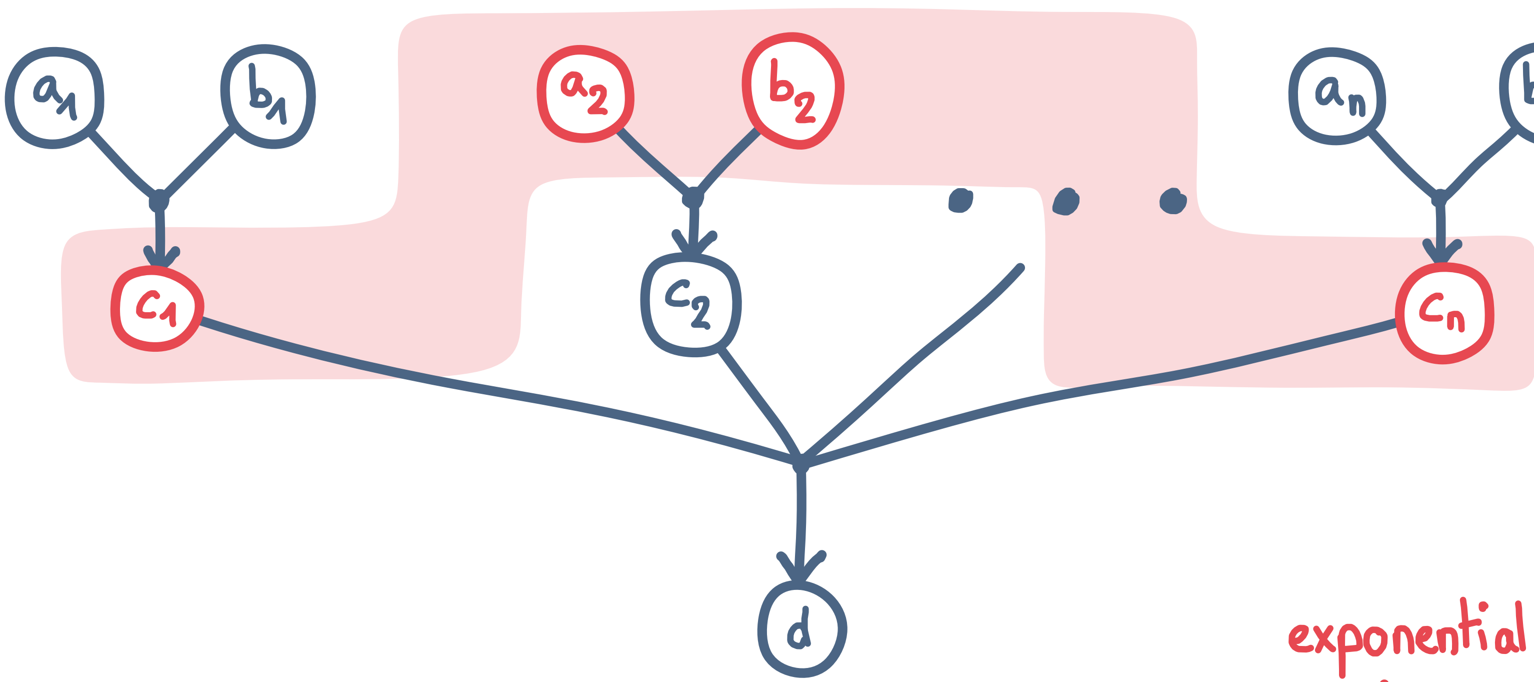
Practical uses :

- seabreeze forecast Adaricheva et al., 23
- stomach cancer risk estimation Nation et al., 21

Exponential blow up

$$X = \{a_1, \dots, a_n, b_1, \dots, b_n, c_1, \dots, c_n, d\}$$

$$\Sigma = \{a_i b_i \to c_i : 1 \leq i \leq n\} \cup \{c_1 \cdots c_n \to d\}$$



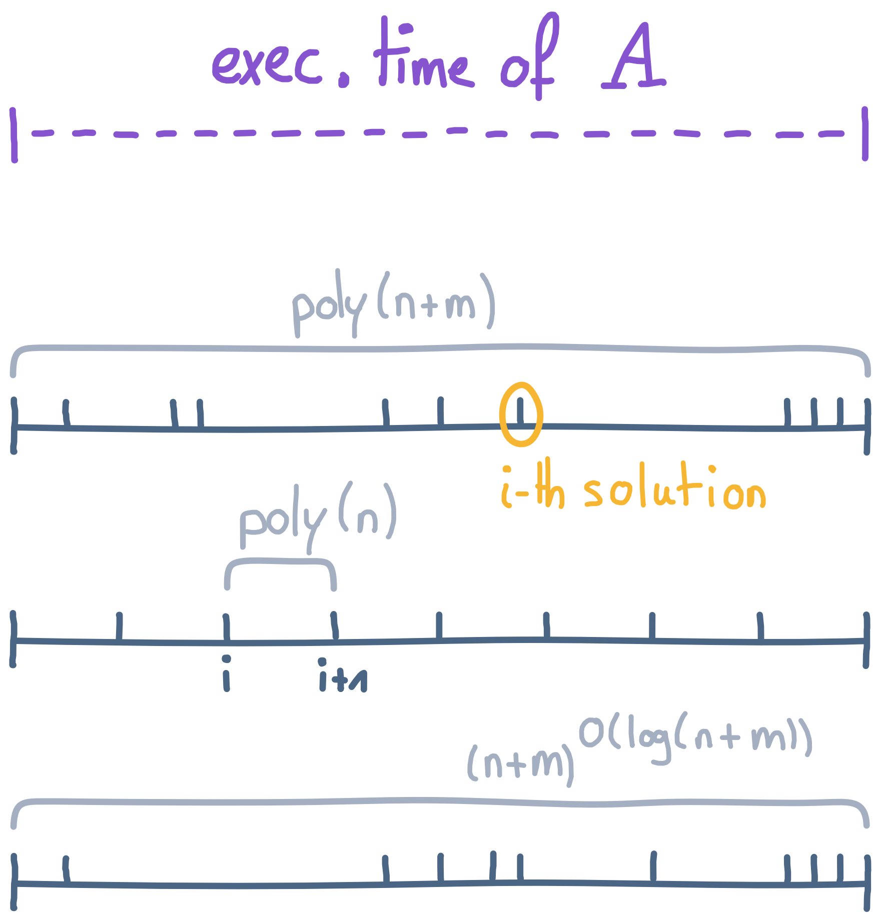
exponential gap w.r.t. Σ!

$$\Sigma_D = \Sigma \cup \boxed{\left\{A \to d : A \in \prod_{i=1}^{n} \{a_i b_i, c_i\}\right\}}$$

Enumeration: output-sensitive complexity

Each of size poly(x)

Enumeration task: with input x, list a set of solutions $R(x)$



Enumeration algorithm A

input x of size n

output $R(x)$ of size m

Output polynomial time

polynomial delay

Output quasi-polynomial time

Related works : part I

PROB : given irreducibles M_i over X,
find the Δ-base Σ_Δ

- algorithm based on Hypergraph dualization Adaricheva, Nation, 17
produces (possibly large) superset of Δ-base

PROB : given an implicational base Σ over X, find the D-base Σ_D

- algorithm using simplification logic <u>Rodriguez et al., 15, 17</u>
  no (output-sensitive) complexity analysis
- poly-delay algorithm listing D-minimal keys <u>Ennaoui, Nourine, 16</u>
  based on solution-graph traversal
  ($\simeq$ D-gen of some x)

Our results: part I

PROB: given irreducibles M_i over X, find the D-base Σ_D

Dualization of distributive lattices + Elbassioni 22

ANV, 24+

THM: given M_i over X, Σ_D can be computed in output quasi-polynomial time

Our results : part II

PROB : given an implicational base Σ over X, find the D-base Σ_D

↓

Solution-graph traversal
\+ Ennaoui, Nourine 16

↓

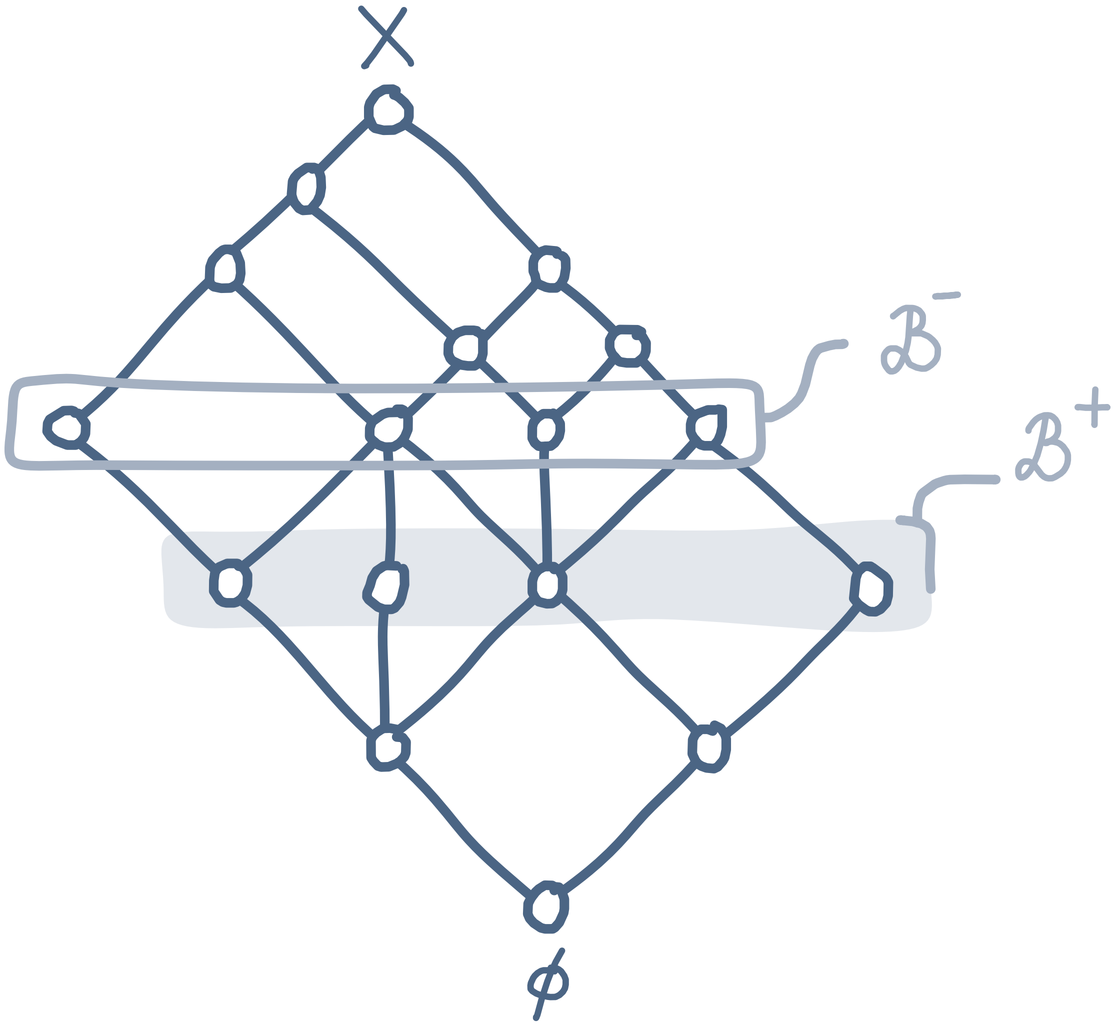
ANV, 24+

THM : given Σ over X, Σ_D can be computed with polynomial delay

PROB: given irreducibles M_i over X,

find the Δ-base Σ_Δ

Dualization

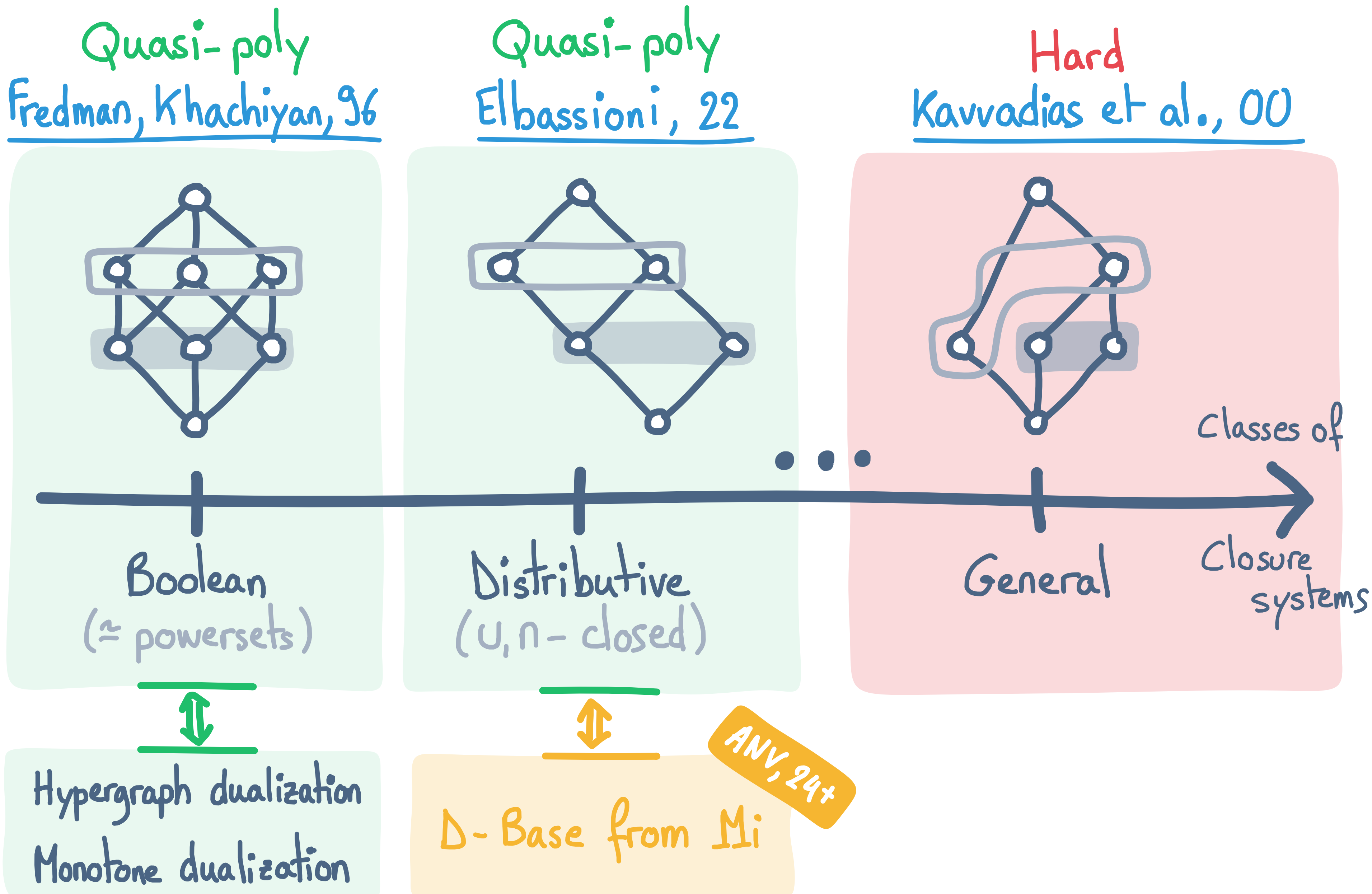


$\mathcal{B}^-$ and $\mathcal{B}^+$ are dual:

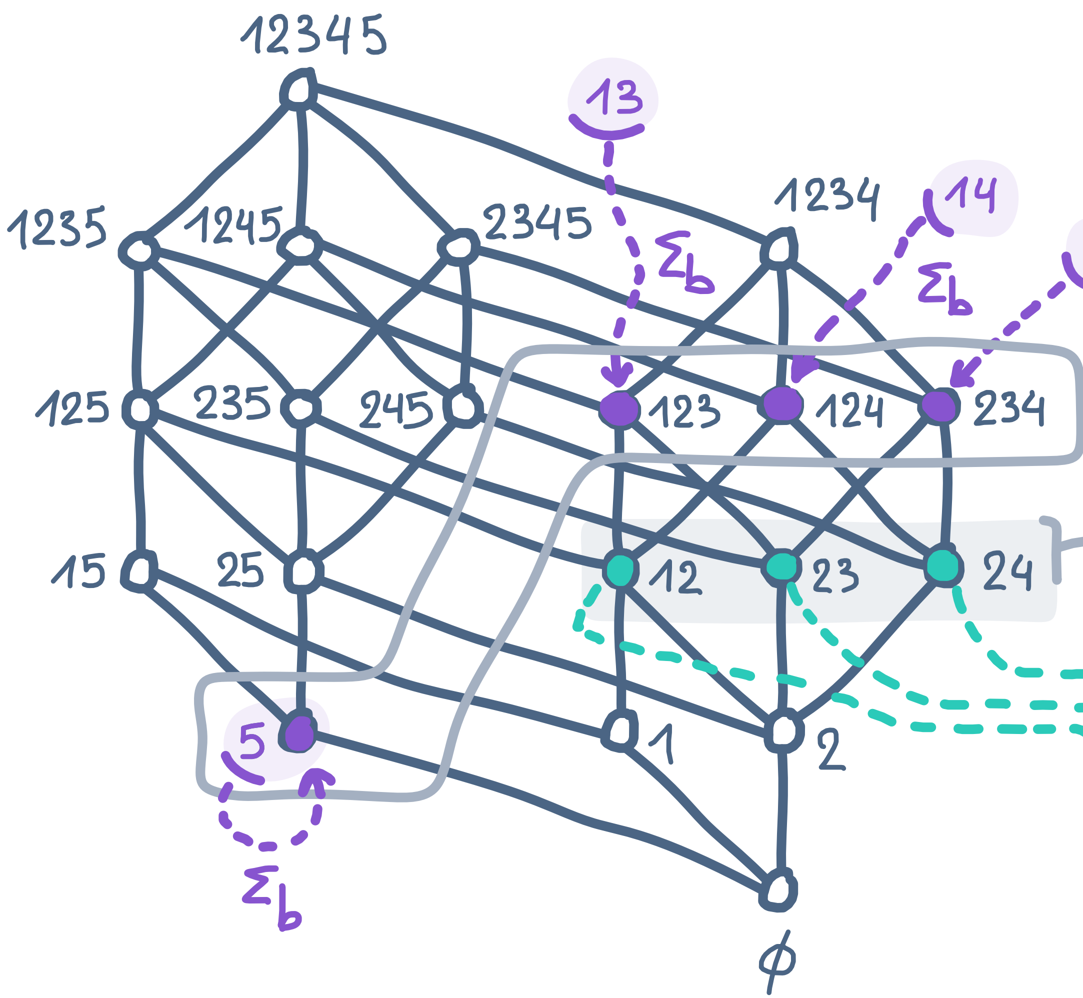
- $\downarrow\mathcal{B}^+ \cup \uparrow\mathcal{B}^- = \mathcal{F}$
- $\downarrow\mathcal{B}^+ \cap \uparrow\mathcal{B}^- = \emptyset$

PROB: with Σ over X and antichain $\mathcal{B}^+$, find antichain $\mathcal{B}^-$

Dualization complexity and D-base

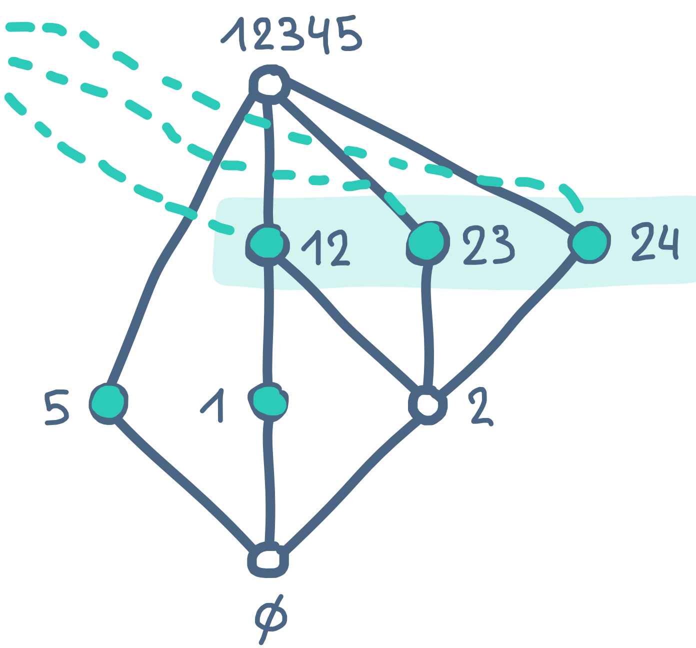


Intuition: Δ-base from Mi & Dualization Distr.



② $\mathcal{B}^-$ obtained from 5 and D-generators of 5

③ $\mathcal{B}^+$ = ⊆-max irreducibles without 5

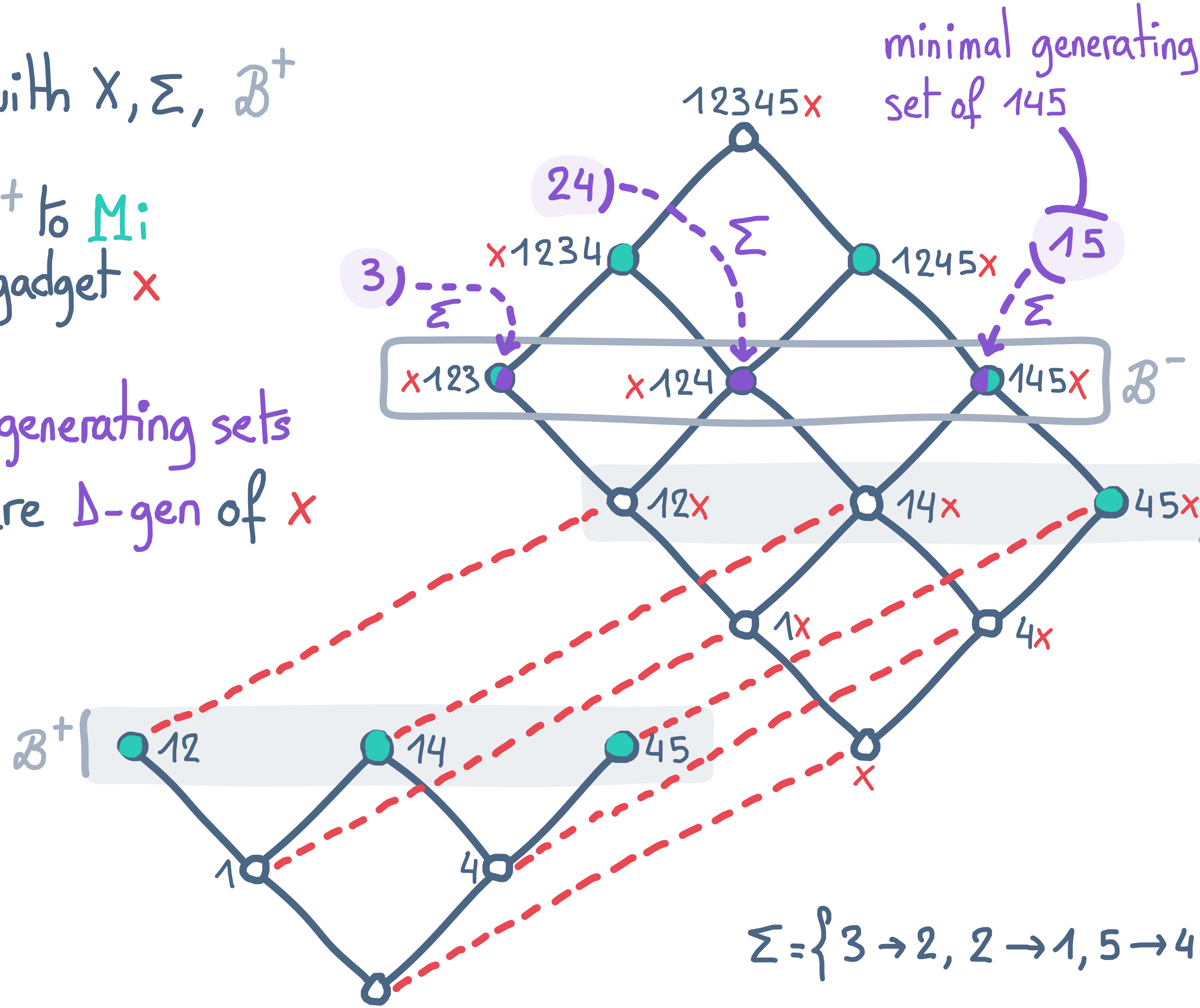


① distributive closure system of

$\Sigma_b = \{3 \to 2,\ 4 \to 2\}$

Intuition: D-base from Mi ≈ Dualization Distr.

1. Start with $X, \Sigma, \mathcal{B}^+$
2. Add $\mathcal{B}^+$ to Mi using gadget x
3. (min.) generating sets of $\mathcal{B}^-$ are D-gen of x



$\Sigma = \{3 \to 2,\ 2 \to 1,\ 5 \to 4\}$

PROB: given irreducibles M_i over X, find the Δ-base Σ_Δ

Dualization of distributive lattices + Elbassioni 22

THM: given M_i over X, Σ_Δ can be computed in output quasi-polynomial time

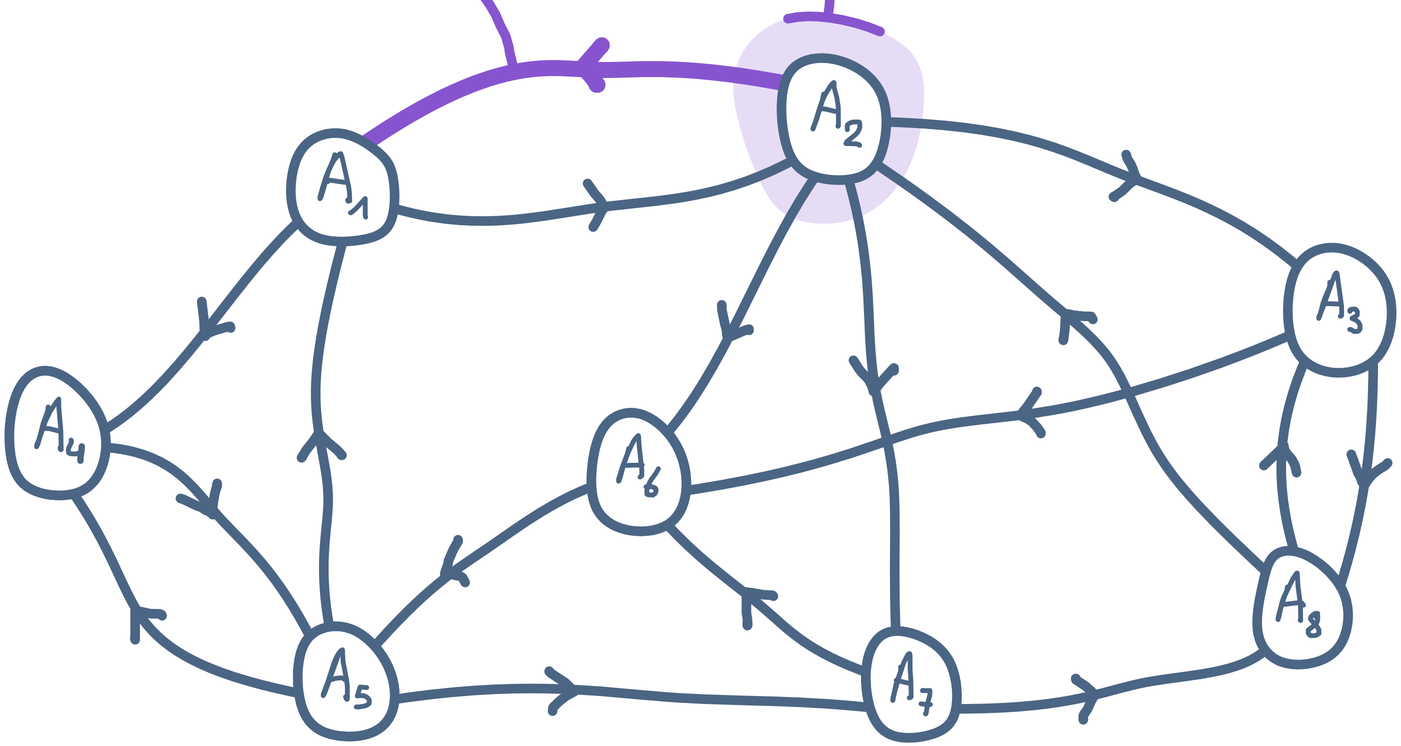
ANV, 24+

PROB: given an implicational base Σ over X, find the D-base Σ_D

Principle: Solution-graph traversal

arc = transition* which turns A_2 into A_1

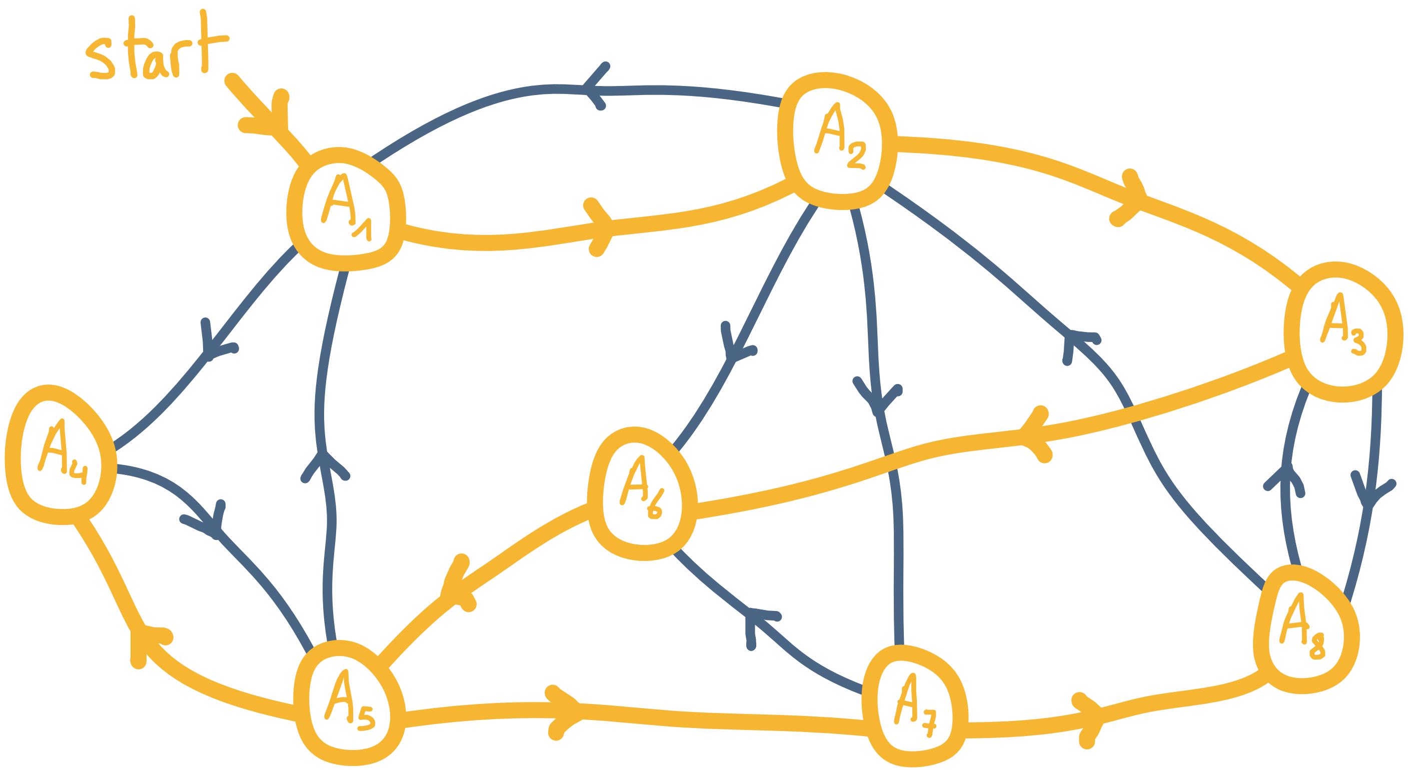
vertex = a D-generator of x



* transition key idea: substitute $a_2 \in A_2$ with B s.t. $B \to a_2 \in \Sigma$

(greedily) minimize w.r.t. Σ_b

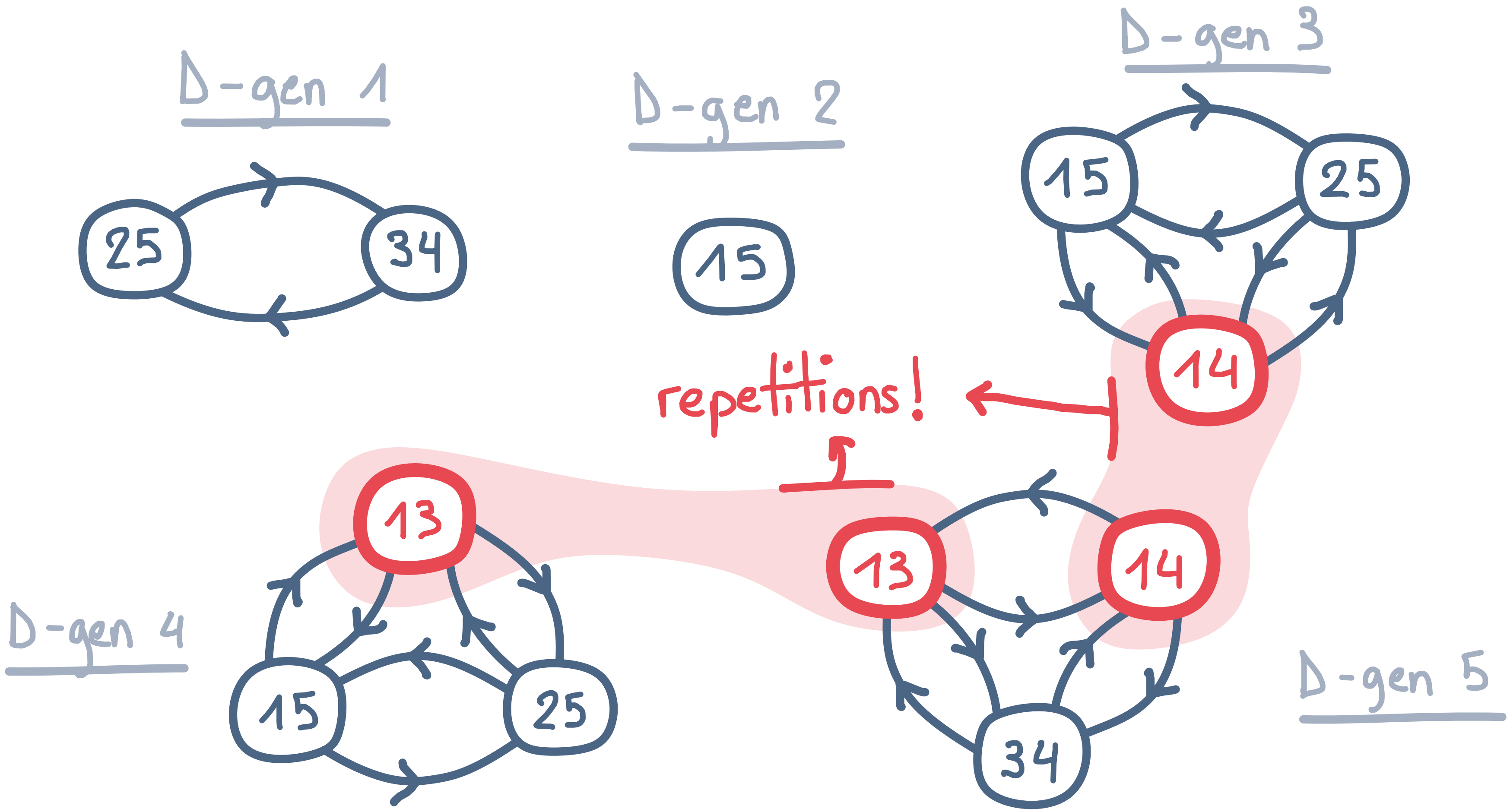
Principle: Solution-graph traversal



1st solution in poly-time + poly transitions + strongly connected

⇒ poly-delay enumeration (with DFS) of D-gen of some x

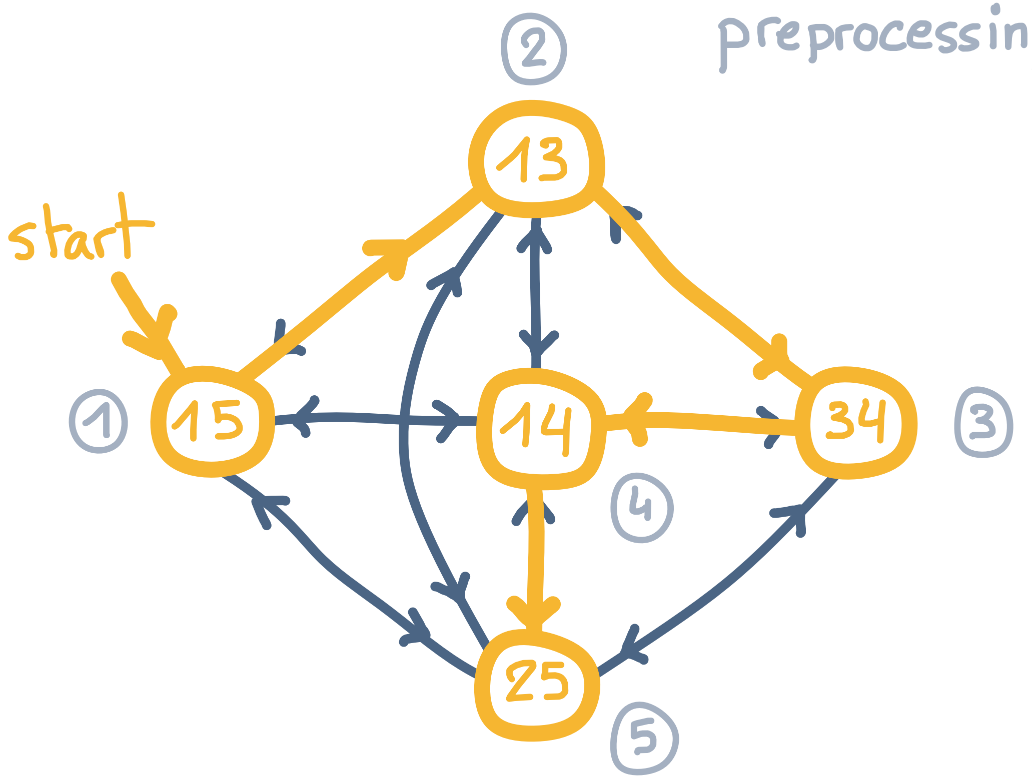
In our case (running ex)



PROB: applying algo on each $x \in X$ yields repetitions

$\Rightarrow$ no guarantee on delay

Fix: merge the graphs



preprocessing

0. 3 → 2, 4 → 2
1. 15 → 2, 15 → 3, 15 → 4
2. 13 → 2, 13 → 5
3. 34 → 5, 34 → 1
4. 14 → 3, 14 → 5
5. 25 → 3, 25 → 1, 25 → 4

FIX: take the union of supergraphs

- poly transitions
- 1st solution in poly-time $\forall x \in X$
- strongly connected components

$\Rightarrow$ poly delay enumeration of all D-gens (with DFSs)

PROB: given an implicational base Σ over X, find the D-base Σ_D

Solution graph traversal

\+ Ennaoui, Nourine 16

THM: given Σ over X, Σ_D can be computed with polynomial delay

ANV, 24+

The D-base:

- describe a lattice by minimal join covers
- ordered direct subset of canonical direct base

Finding the D-base:

- output quasi-poly from $\mathcal{M}_i$
- poly-delay from Σ

Questions regarding E-base (subset of D-base)

- Characterize systems with valid E-base
- Similar algorithms for E-base?

References

Dilworth, 40

Dilworth
Lattices with unique irreducible decompositions
Annals of Mathematics, 1940

Finkbeiner, 51

Finkbeiner
A general dependence relation for lattices
Proceedings of the AMS, 1951

Jónsson, Nation, 77

Jónsson, Nation
A report on sublattices of a free lattice
Contributions to Universal Algebra, 1977

References

Gaskill, Rival

Gaskill, Rival, 78

An exchange property for modular lattices

Algebra Universalis, 1978

Day

Day, 79

Characterizations of finite lattices that are bounded homomorphic images or sublattices of free lattices

Canadian Journal of Mathematics

Faigle

Faigle, 86

Exchange properties of combinatorial closure spaces

Discrete Applied Mathematics, 1986

References

Nation — Nation, 90

An approach to lattice varietes of finite height

Algebra Universalis, 1990

Freese, Jesek, Nation — Freese et al., 95

Free lattices

Survey and Monograph of the AMS, 1995

Bertet, Monjardet — Bertet, Monjardet, 10

The multiple facets of the canonical direct unit implicational basis

Theoretical Computer Science, 2010

References

Adaricheva, Nation, Rand — Adaricheva et al., 13
Ordered direct implicational basis of finite closure system
Discrete Applied Mathematics

Rodriguez-Lorenzo, Adaricheva, Cordero, Enciso, Mora — Rodriguez et al, 17
Formation of the D-basis from implicational system using Simplification Logic
International Journal on General Systems, 2017

Rodriguez-Lorenzo, Adaricheva, Cordero, Enciso, Mora — Rodriguez et al, 15
From an implicational system to its corresponding D-basis
2015

References

Adaricheva, Nation — Adaricheva, Nation, 17

Discovery of the D-basis in binary table based on hypergraph dualization

Theoretical Computer Science, 2017

Adaricheva, Nourine, Vilmin — Adaricheva, Nourine, Vilmin, 24+

Computing the D-base and D-relation of finite closure systems

arXiv, submitted to Theoretical Computer Science

Adaricheva, Bernhardt, Liu, Schmidt — Adaricheva et al., 23

Importance of overnight parameters to predict sea breeze on Long Island

2023

References

Nation, Cabot-Miller, Segal, Lucito, Adaricheva — Nation et al., 21

Combining algorithms to find signatures that predict risk in early stage of stomach cancer

Journal of Computational Biology, 2021

Ennaoui, Nourine — Ennaoui, Nourine, 16

Polynomial delay hybrid algorithms to enumerate candidate keys for a relation

BDA, 2016

Fredman, Khachiyan — Fredman, Khachiyan, 96

On the complexity of dualization of monotone disjunctive normal forms

Journal of Algorithms, 1996

References

Elbassioni — Elbassioni, 22

On dualization over distributive lattices

Discrete Mathematics and Theoretical Computer Science, 2022

Kavvadias, Sideri, Stavropoulos — Kavvadias et al., 00

Generating maximal models of a Boolean expression

Information Processing Letters, 2000