THE D-BASE OF FINITE CLOSURE SYSTEMS

Simon Vilmin (LIS, Aix-Marseille Université, France)

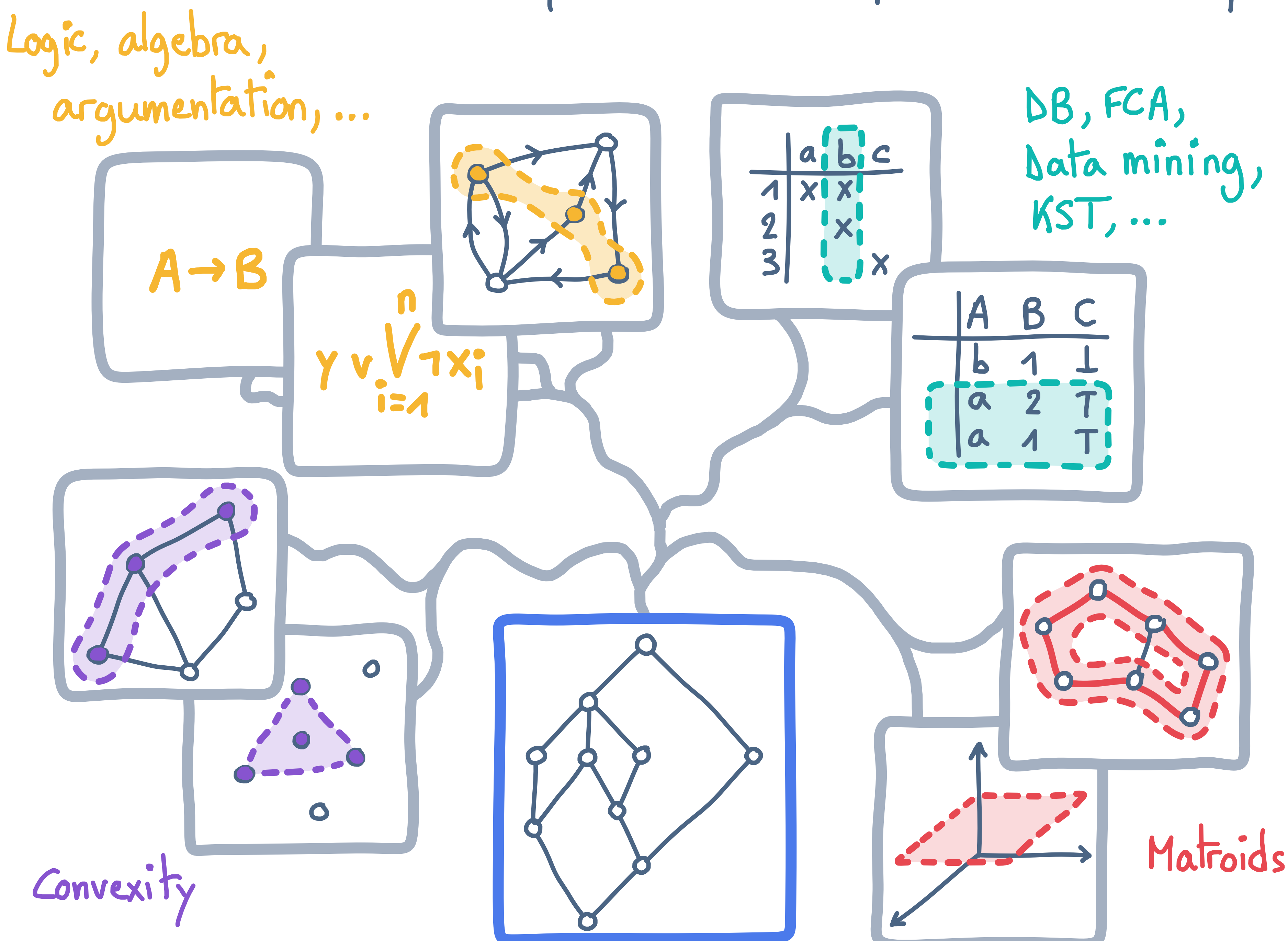
NY Combinatorics Seminar - March, 7th 2025

Joint work with:

Kira Adaricheva (Math dpt., Hofstra University)
Lhouari Nourine (LIMOS, Université Clermont Auvergne, France)

(arXiv link)

Closure systems (and their representations) are ubiquitous



Outline

PART I: what is the D-base

from closure systems …

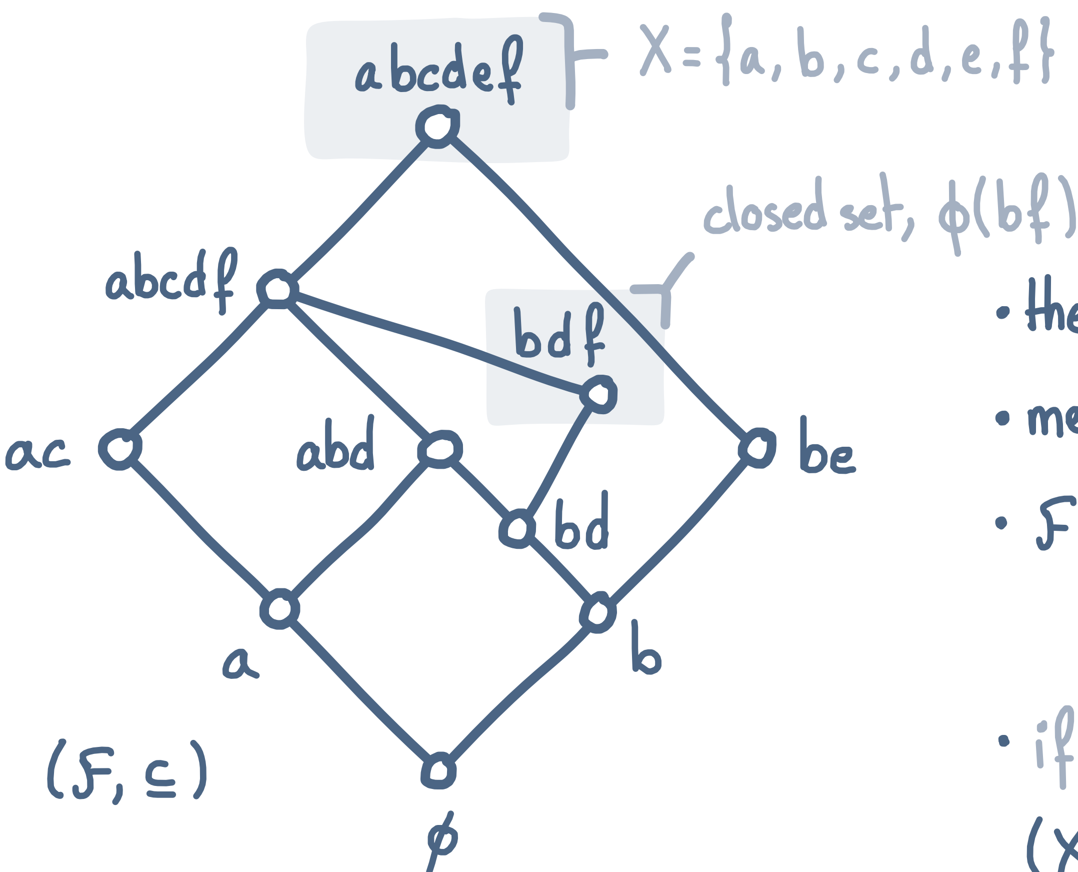
… to implications

PART II: computing the D-base from irreducible closed sets

PART I : What is the D-base

First : what is a closure system ?

DEF: a closure system is a pair $(X, \mathcal{F})$ where X is a groundset and $\mathcal{F} \subseteq 2^X$ is closed under intersection and contains X.



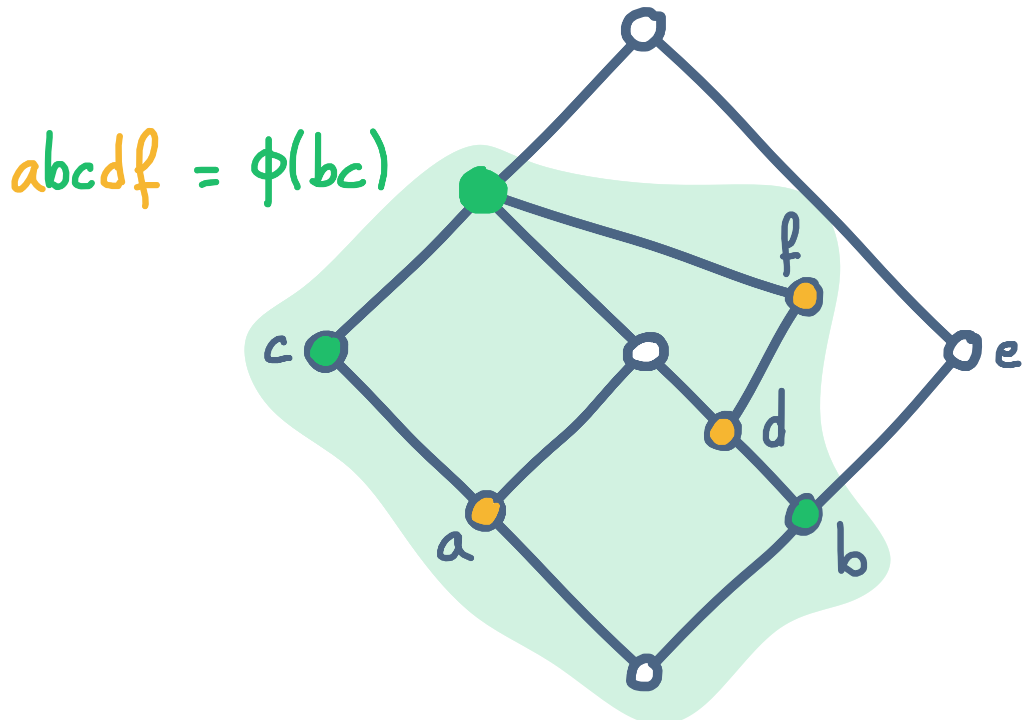
$(\mathcal{F}, \subseteq)$

- the poset $(\mathcal{F}, \subseteq)$ is a (closure) lattice
- members of $\mathcal{F}$ are closed sets
- $\mathcal{F}$ induces a closure operator ϕ

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

- if $\mathcal{F}$ also closed under union, $(X, \mathcal{F})$ is distributive

Reconstructing $\mathcal{F}$ from X and the diagram



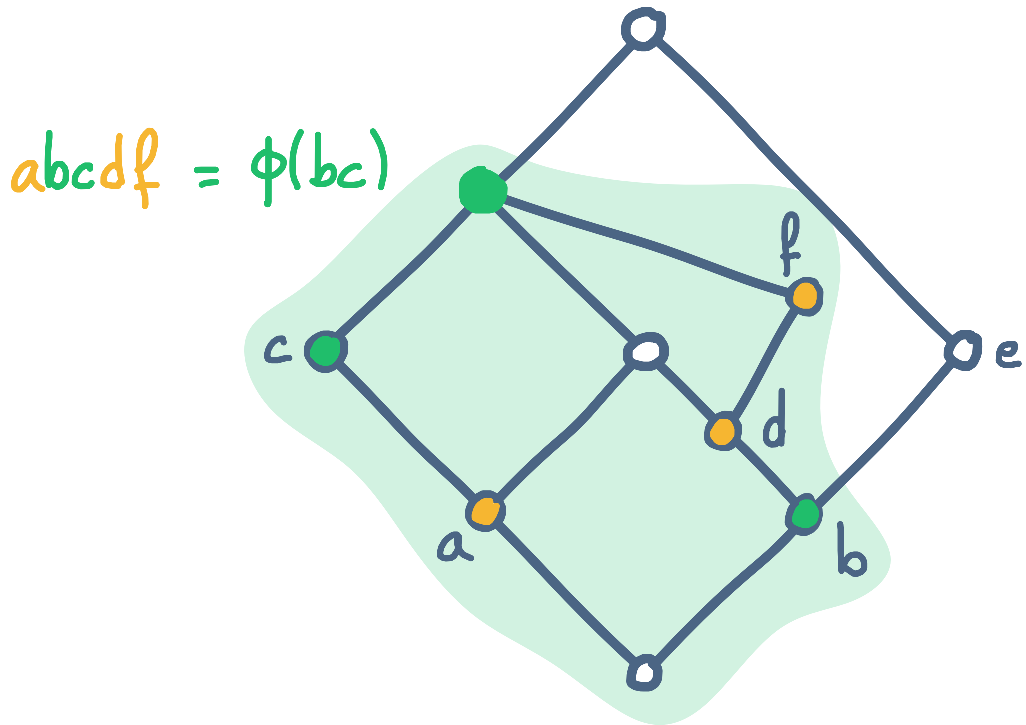
How to (visually) compute closures with the diagram:

(1) for each $x \in X$, give the label x to $\phi(x)$

(2) for $A \subseteq X$, take the minimal point above the labels of A

(3) $\phi(A)$ is the set of labels below this point

A relation on X to describe F



$\phi(bc)$ captures a, d, f. We can write this as implications

$bc \to a$, $bc \to d$, $bc \to f$

IDEA: describe the (structure of the) closure system by means of implications

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 (implications of the 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)

Freese et al., 95 : book on free lattices

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

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

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

our contribution

Adaricheva, Nourine, Vilmin, 25+ : output-sensitive study of Δ-base computation, with implications or irreducible closed sets (= binary data)

IDEA: describe the (structure of the) closure system by means of implications

is the same as finding, e.g.:

- a Horn CNF for a Horn function (logic)
- a cover of Functional Dependencies in a relation (databases)
- association rules (with support 1) in transactions (data mining)

and generalizes finding, e.g.:

- the circuits of a matroid
- the minimal transversals of a (hyper)graph

DEF: an *implication* over X is a statement $A \to x$, with $A \subseteq X$ and $x \in X$. An *implicational base (IB)* is a pair (X, Σ) where Σ is a *set of implications* over X.

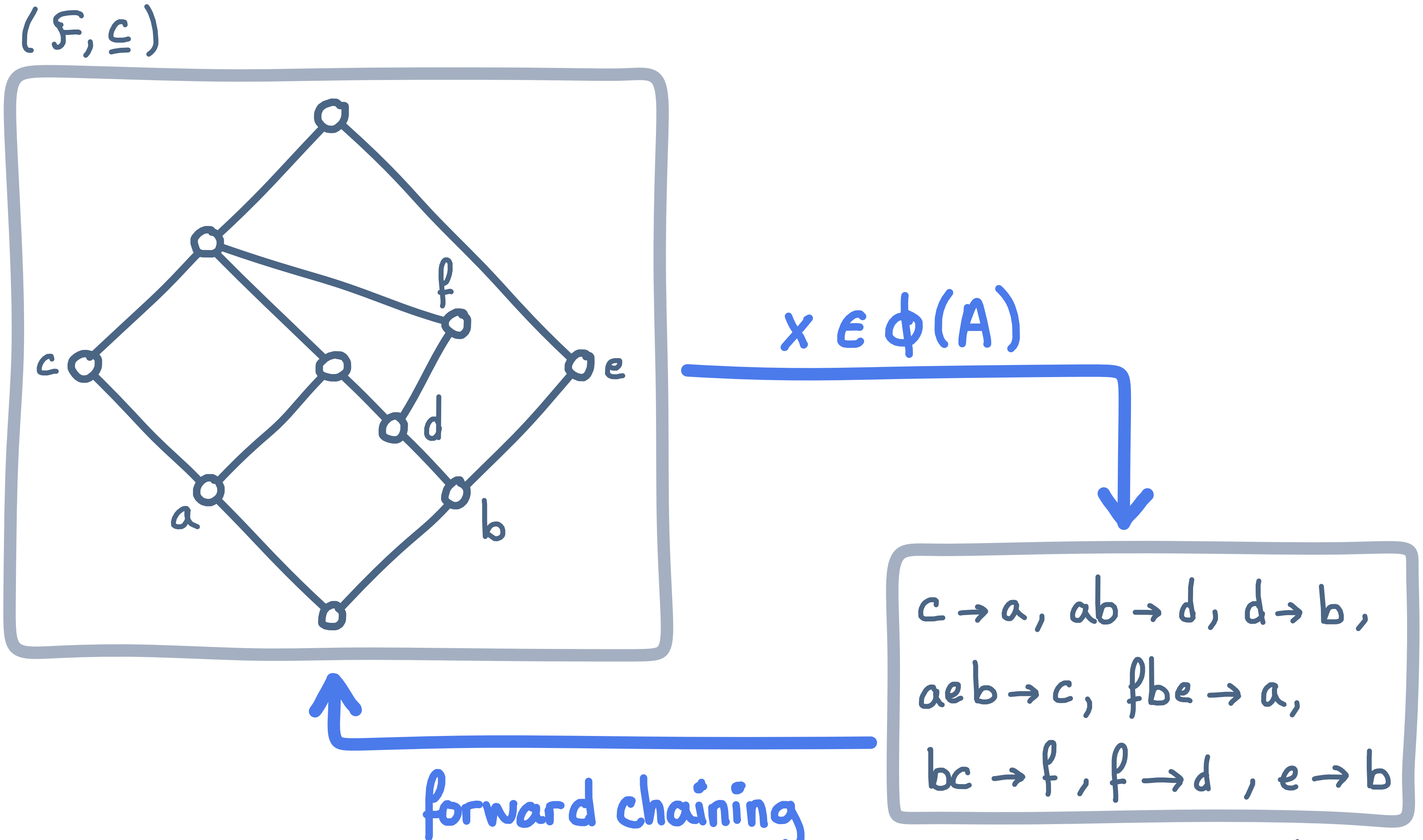
An IB (X, Σ) encodes a closure system $(X, \mathcal{F})$ where:

$$\mathcal{F} = \{ F : A \subseteq F \text{ implies } x \in F \ \forall A \to x \in \Sigma \}$$

A closure system can be represented by many (equivalent) IBs

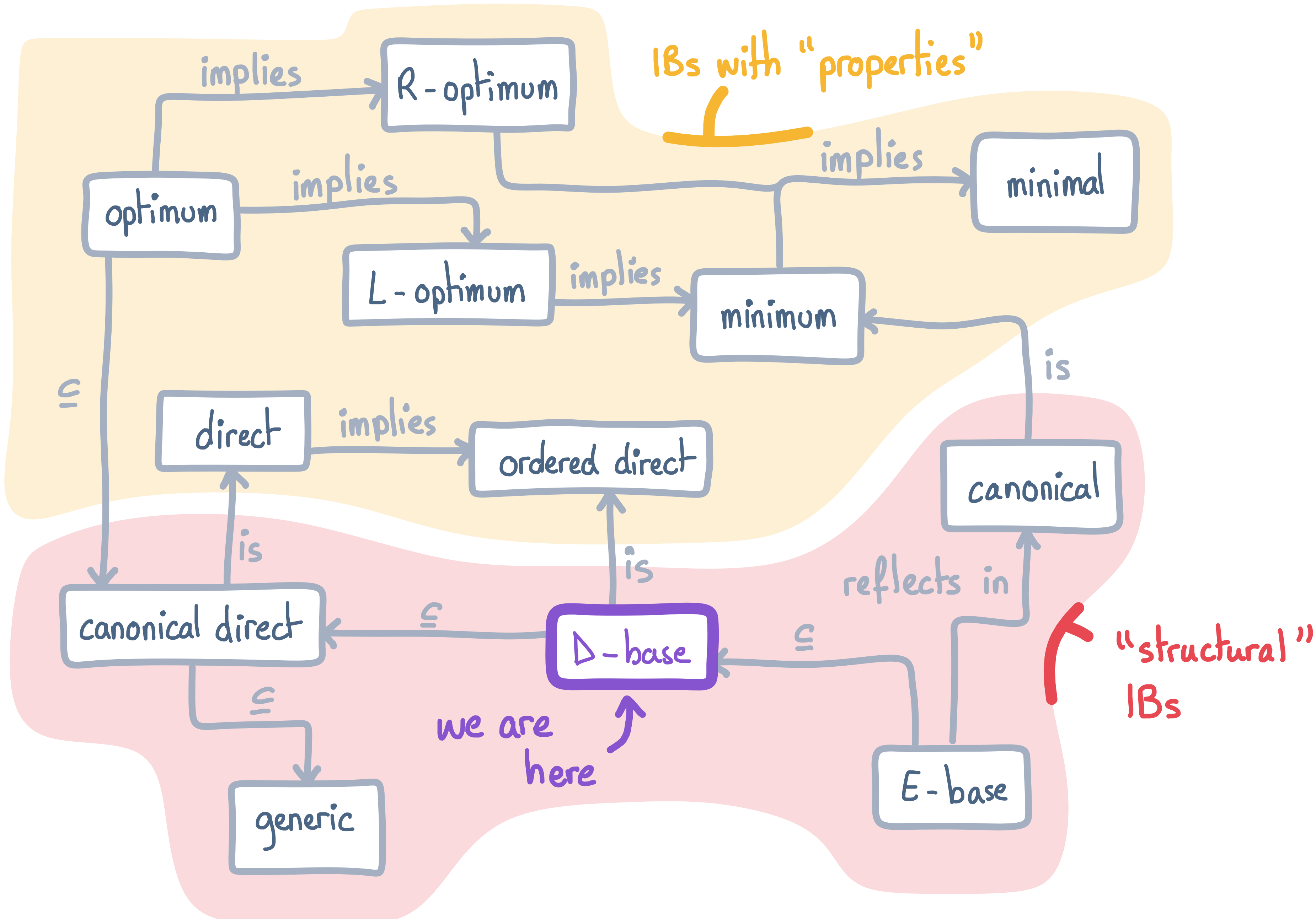
$A \to x$ holds in $(X, \mathcal{F})$ iff $x \in \phi(A)$

On our example



An IB (X, Σ)

the (partial) landscape of IBs

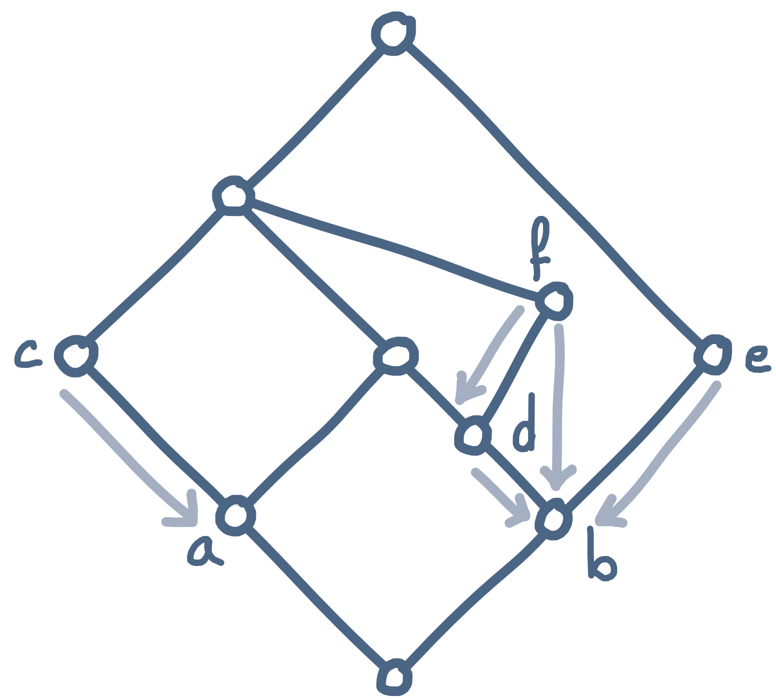


IDEA : for each $x \in X$, find all the "minimal explanations" of x

The way elements of X are ordered :

a is below c $(a \in \phi(c))$

d is below f $(d \in \phi(f))$



Represented by binary implications :

$c \to a$, $f \to d$, $f \to b$, $d \to b$, $e \to b$

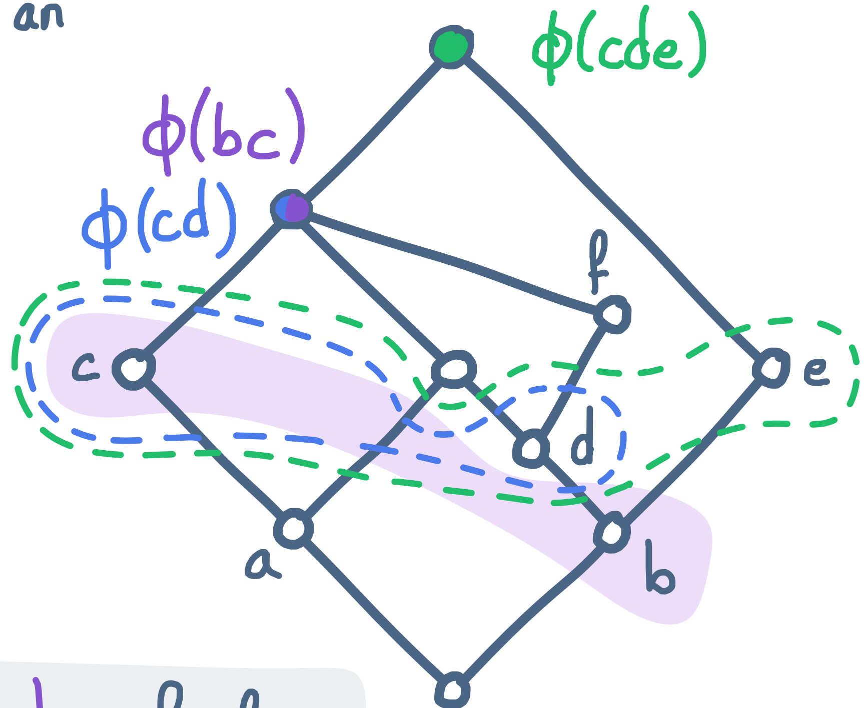
IDEA : for each $x \in X$, find all the "minimal explanations" of x

Minimal (non-singleton) sets capturing an element

$f \in \phi(cde)$

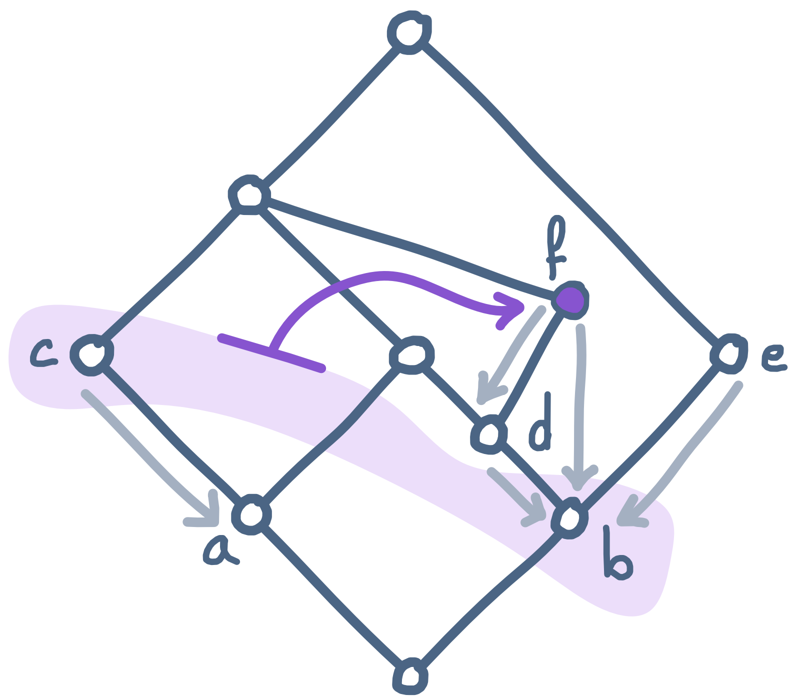
$f \in \phi(cd)$ and $cd \subset cde$

$f \in \phi(bc)$ and bc "refines" cd



bc is minimal: it is a Δ-generator of f.
The implication $bc \to f$ will be in the Δ-base

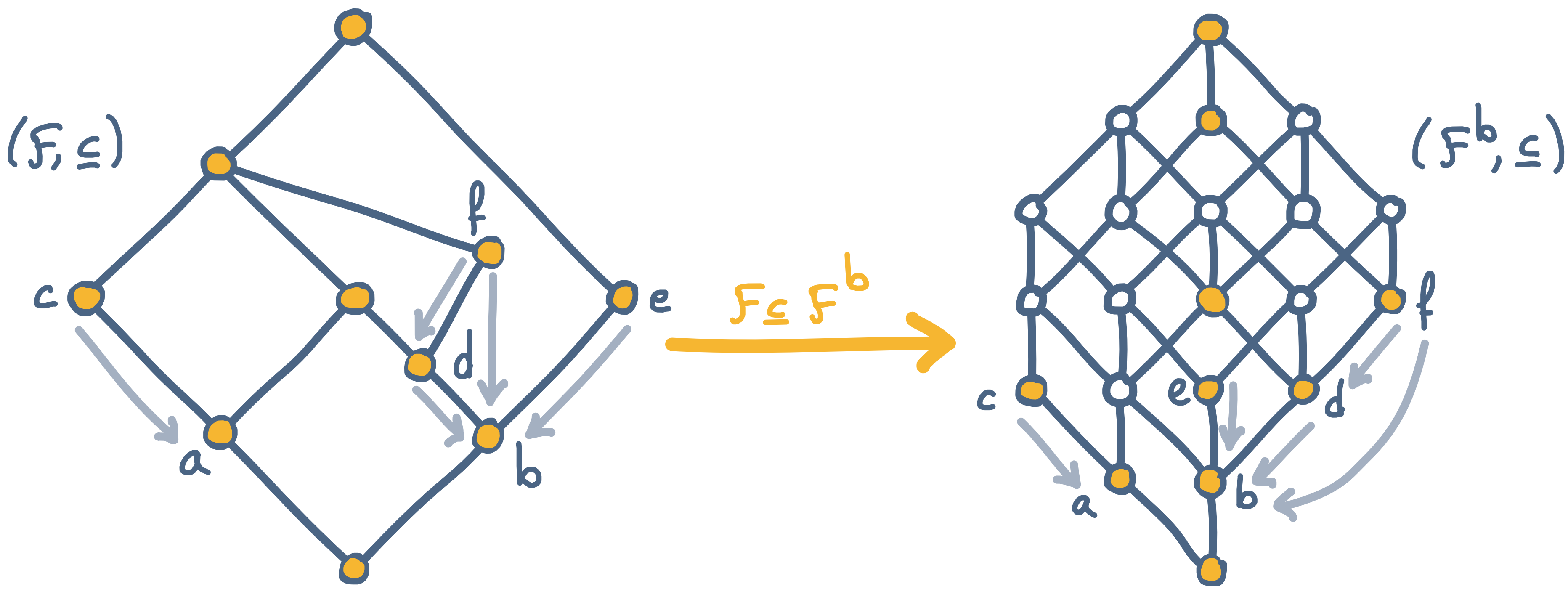
IDEA : for each $x \in X$, find all the "minimal explanations" of x



Δ-base = order on X + Δ-generators

Formalization : binary part

DEF: the binary part of $(X, \mathcal{F})$ is the closure system $(X, \mathcal{F}^b)$ with $\phi^b(A) = \bigcup \{\phi(a) : a \in A\}$ for $A \subseteq X$. The IB (X, Σ^b) with $\Sigma^b = \{a \to x : x \in \phi(a), x \neq a\}$ represents $(X, \mathcal{F}^b)$



$$\Sigma^b = \{c \to a, f \to d, f \to b, d \to b, e \to b\}$$

Formalization: D-generator, D-base Adaricheva et al., 13

DEF: a subset A of X is a D-generator of x if $x \in \phi(A)$ but $x \notin \phi^b(A)$ and $\forall B \subseteq X$:

- $\phi^b(B) = \phi^b(A)$ implies $A \subseteq B$
- $\phi^b(B) \subset \phi^b(A)$ implies $x \notin \phi(B)$

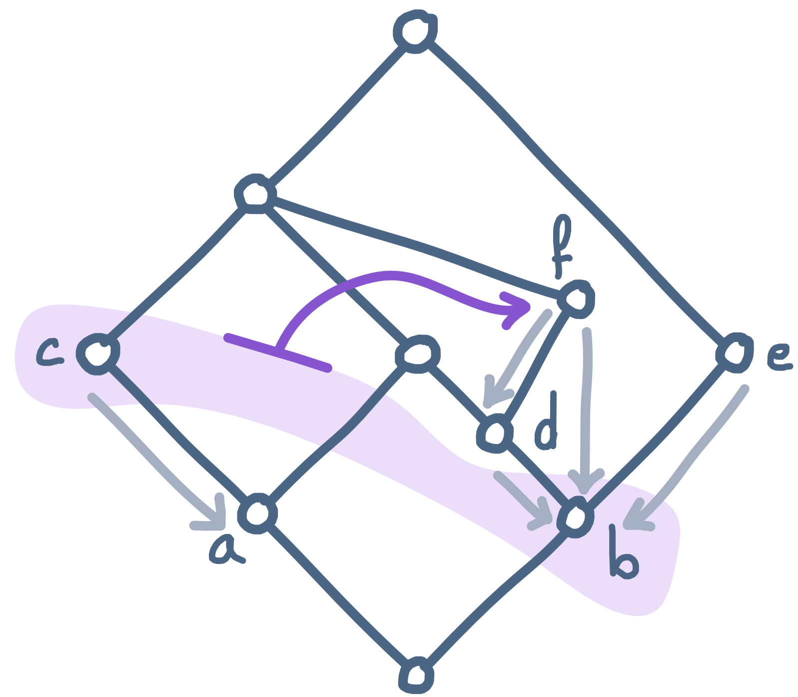
encodes minimality

DEF: the D-base of $(X, \mathcal{F})$ is the IB (X, Σ_D) with

$$\Sigma_D = \Sigma^b \cup \{A \to x : x \in X,\ A \text{ D-gen of } x\}$$

RMK: why is it a valid IB? Take $Y \subseteq X$, any $x \in \phi(Y) \setminus Y$ has a minimal explanation w.r.t. Y: follow Σ^b, find these explanations in $\Sigma_D \setminus \Sigma^b$ and use them.

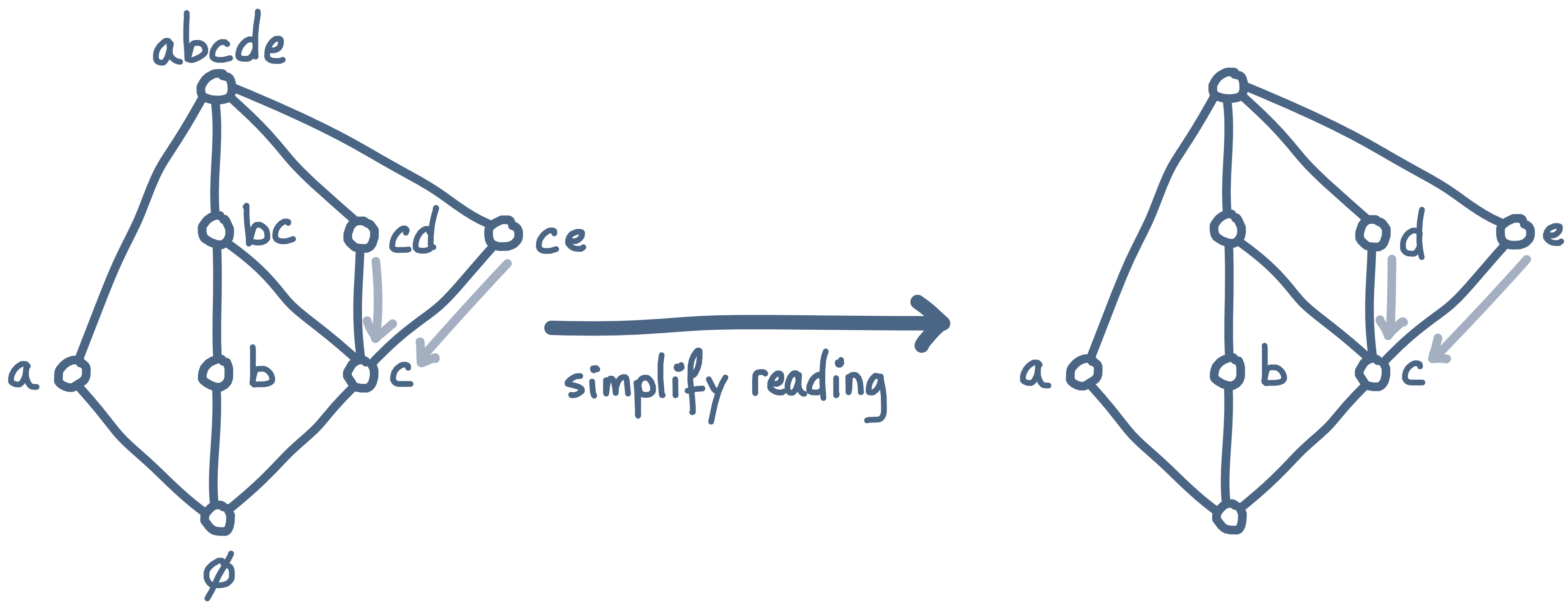
The D-base of our example



$$\Sigma_D = \left[c \to a,\ f \to d,\ f \to b,\ d \to b,\ e \to b \right] \cup \left[de \to a,\ de \to c,\ ae \to c,\ af \to c,\ ab \to d,\ bc \to f,\ de \to f,\ ae \to f \right]$$

PART II: Computing the Δ-base from irreducible closed sets

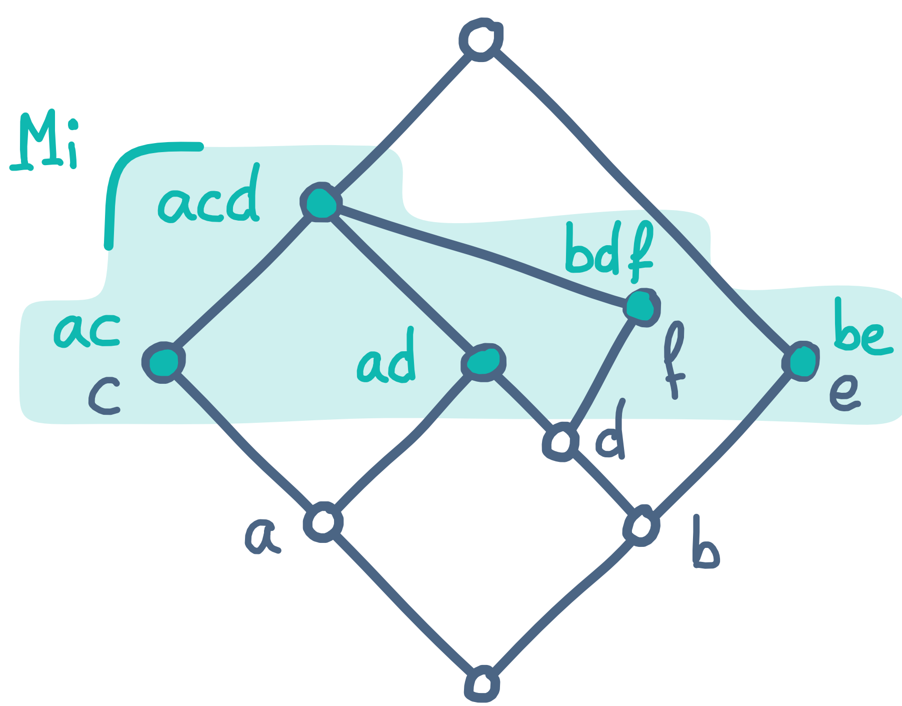
A new example

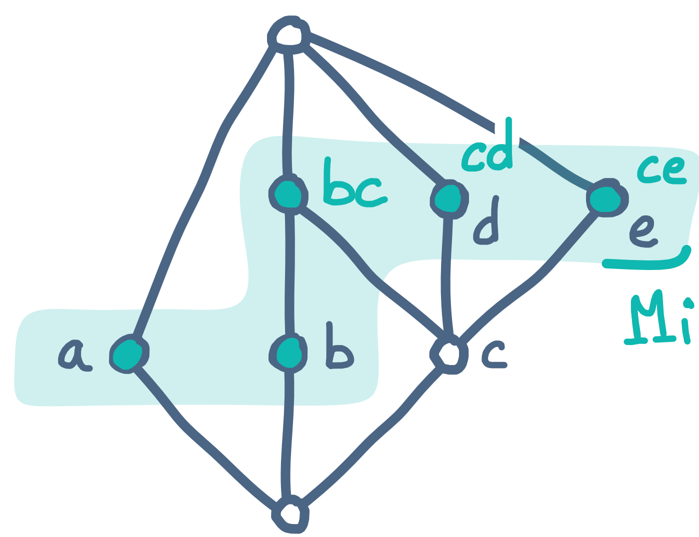


$$\Sigma_\Delta = \{d \to c, e \to c\} \cup \begin{bmatrix} bd \to a,\ be \to a,\ de \to a, \\ ac \to b,\ de \to b, \\ ab \to c, \\ ab \to d,\ ac \to d, \\ ab \to e,\ ac \to e \end{bmatrix}$$

DEF : a closed set M distinct from X is irreducible if it is not the intersection of other closed sets. We put

$$\mathcal{M}_i = \{ M : M \in \mathcal{F}, M \text{ is irreducible} \}$$



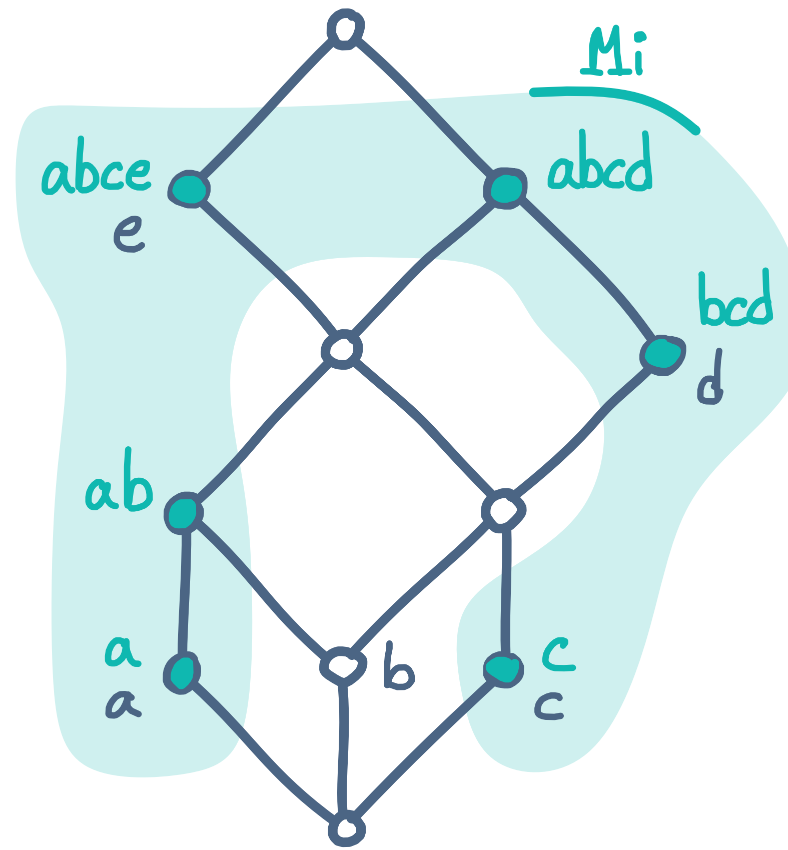


RMK : $\mathcal{F}$ can be built from $\mathcal{M}_i$ using intersections

Irreducibles and binary data

	a	b	c	d	e	
Folavril	x					a
Lazuli			x			c
Dupont	x	x				ab
Chloé		x	x	x		bcd
Wolf	x	x	x	x		abcd
Lil	x	x	x		x	abce

Lil bought product c



ab → c means "any person buying a and b also buys c"

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

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

RMK : enumeration tasks, listing implications without repetitions

Further motivations for computing (X, Σ_D)

Theoretical / algorithmic properties :

- convey structural information of closure systems
- ordered direct (fast forward chaining)

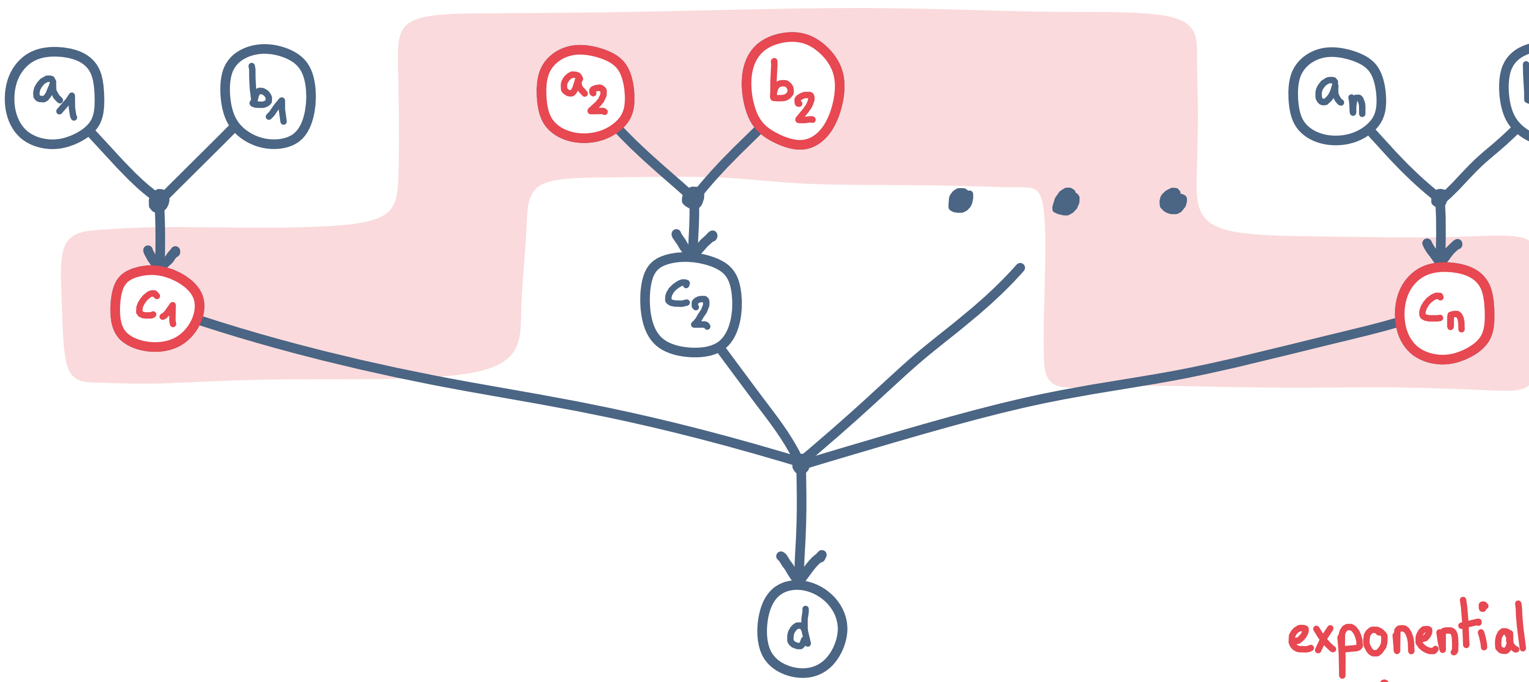
Practical uses :

- 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!

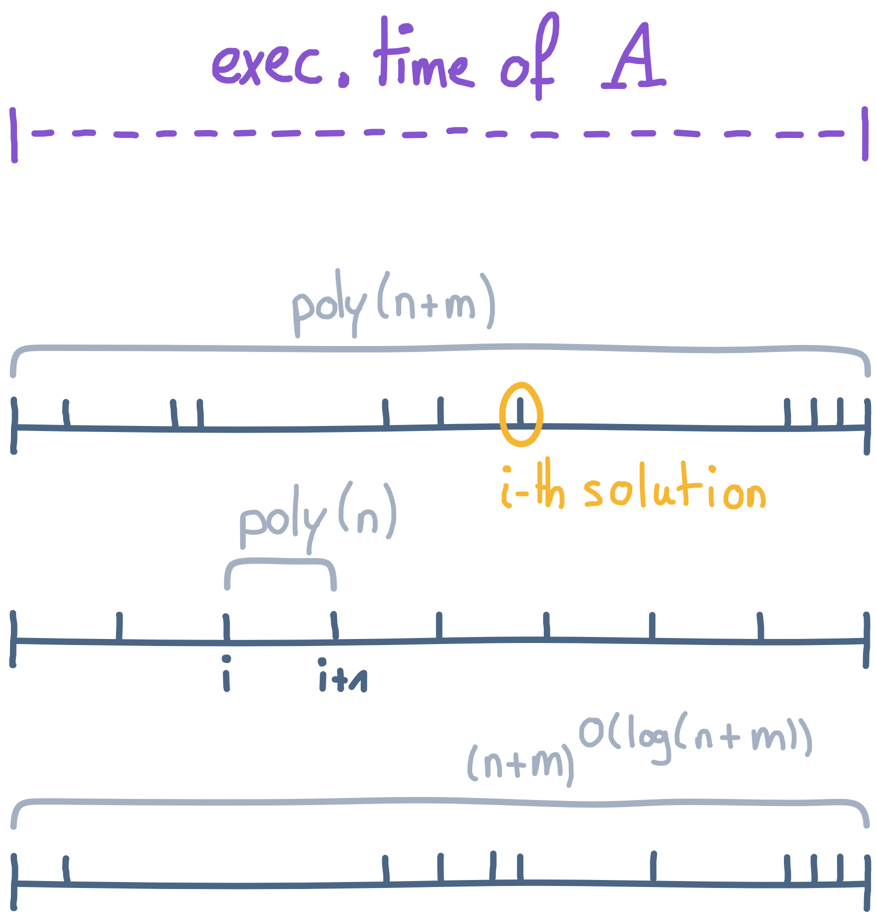
$$\Sigma_\Delta = \Sigma \cup \boxed{\left\{A \to d : A \in \prod_{i=1}^{n} \{a_i b_i, c_i\}\right\}}$$

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

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

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

Related works : part II

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

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

Our results: part I

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

Dualization of distributive lattices + Elbassioni 22

ANV, 25+

THM: given M_i over X, Σ_Δ 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

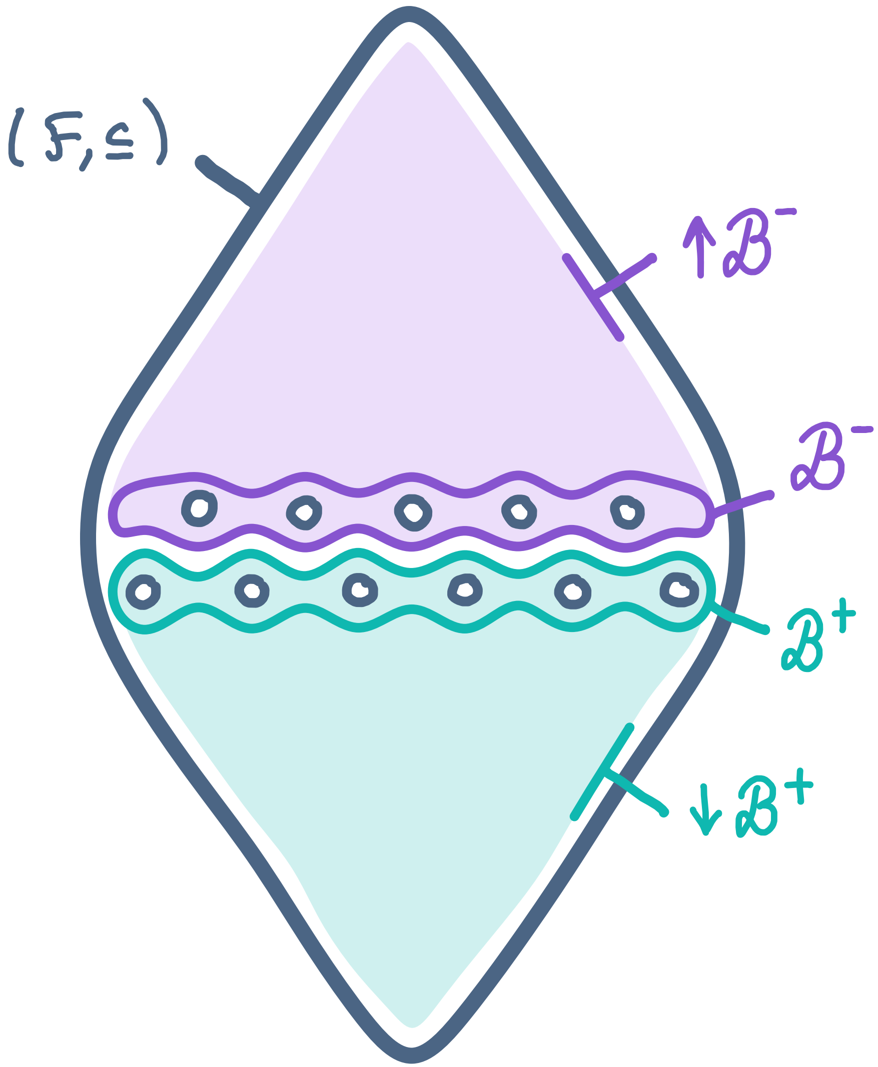
Solution-graph traversal
+ Ennaoui, Nourine 16

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

ANV, 25+

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

Dualization (with IBs)



$\mathcal{B}^-$ and $\mathcal{B}^+$ antichains :

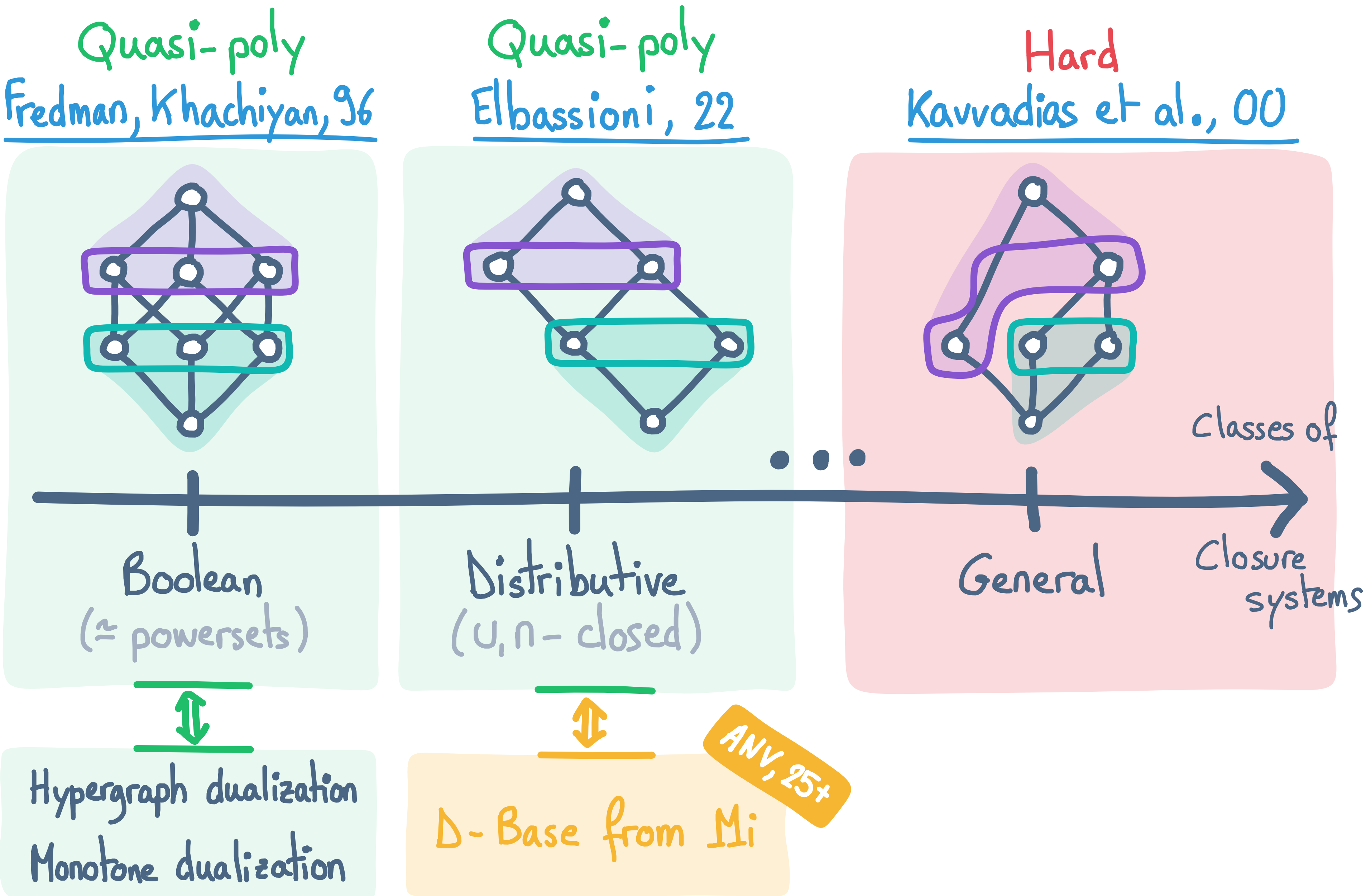
- families of incomparable closed sets

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

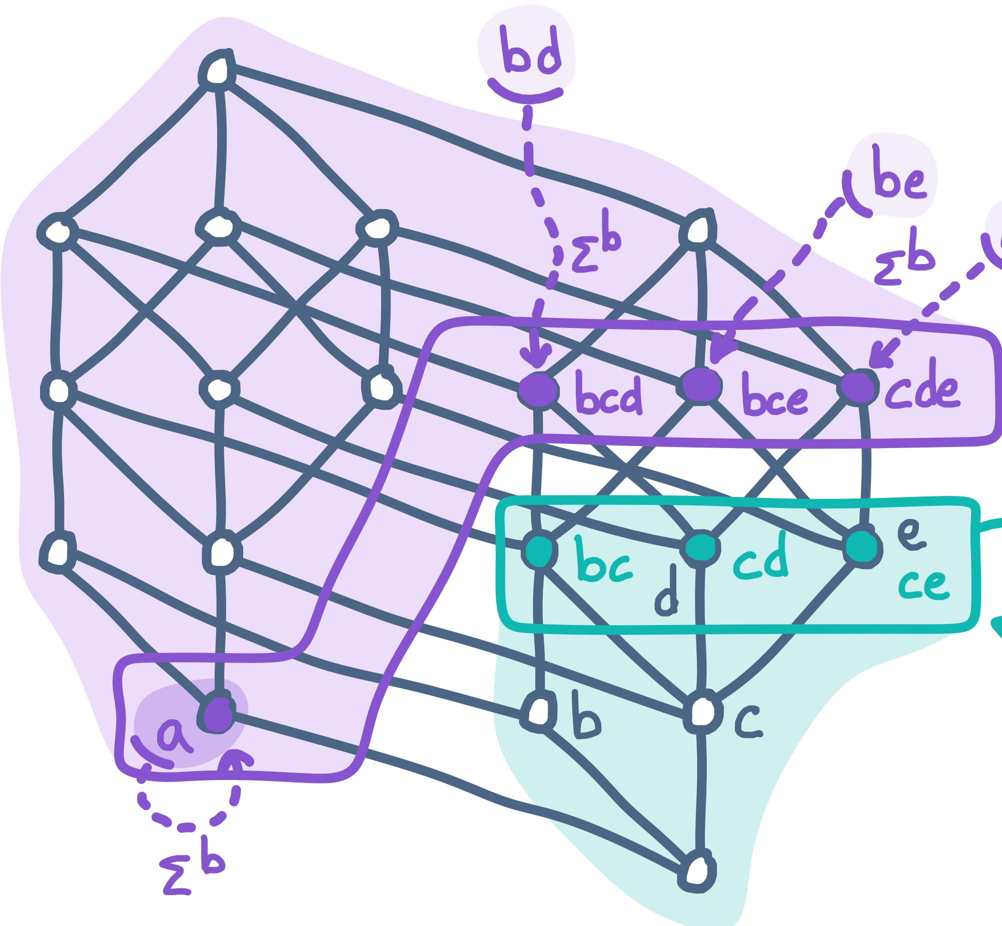
- $\downarrow \mathcal{B}^+ \cup \uparrow \mathcal{B}^- = \mathcal{F}$
- $\downarrow \mathcal{B}^+ \cap \uparrow \mathcal{B}^- = \emptyset$

PROB : given (X, Σ) and $\mathcal{B}^+$, find $\mathcal{B}^-$

Dualization complexity and Δ-base

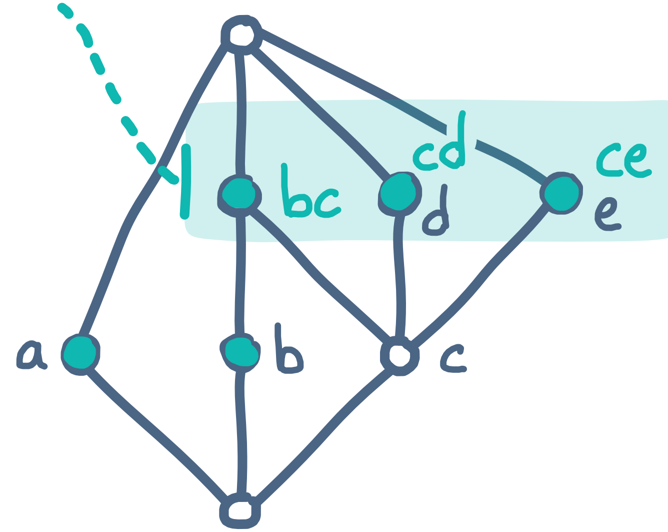


Intuition : D-base from Mi $\leq$ Dual Distributive



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

③ $\mathcal{B}^+ = \subseteq$-max irreducibles without a



① $(\mathcal{F}^b, \subseteq)$ distributive

$\Sigma^b = \{d \to c, e \to c\}$

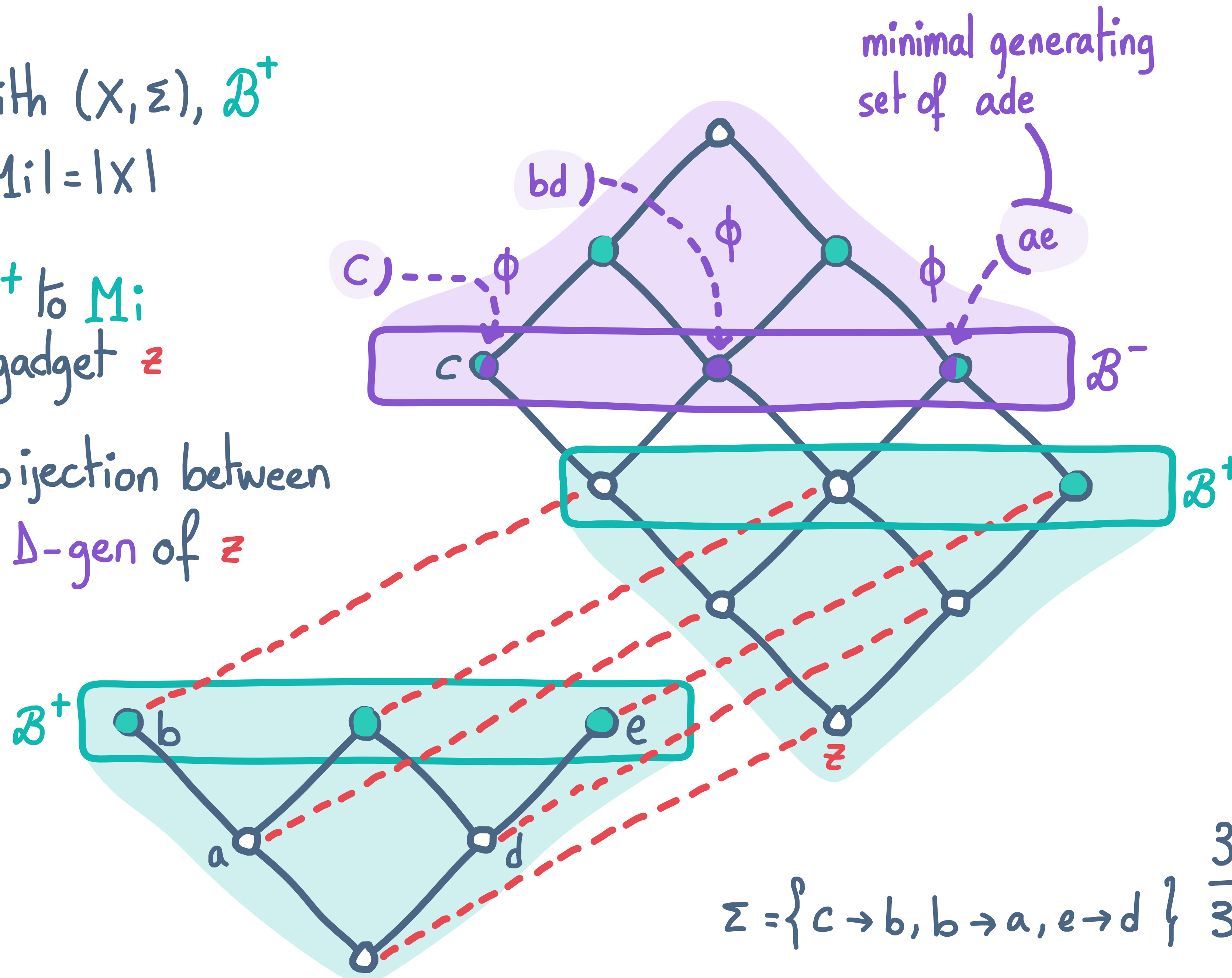
Intuition: Δ-base from Mi $\Rightarrow$ Dual distributive

① Start with (X, Σ), $\mathcal{B}^+$

Distr $\Rightarrow$ $|\text{Mi}| = |X|$

② Add $\mathcal{B}^+$ to Mi using gadget z

③ ϕ is a bijection between $\mathcal{B}^-$ and Δ-gen of z



$\Sigma = \{c \to b, b \to a, e \to d\}$

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

Dualization of distributive lattices + Elbassioni 22

ANV, 25+

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

Conclusion

The D-base:

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

Finding the D-base:

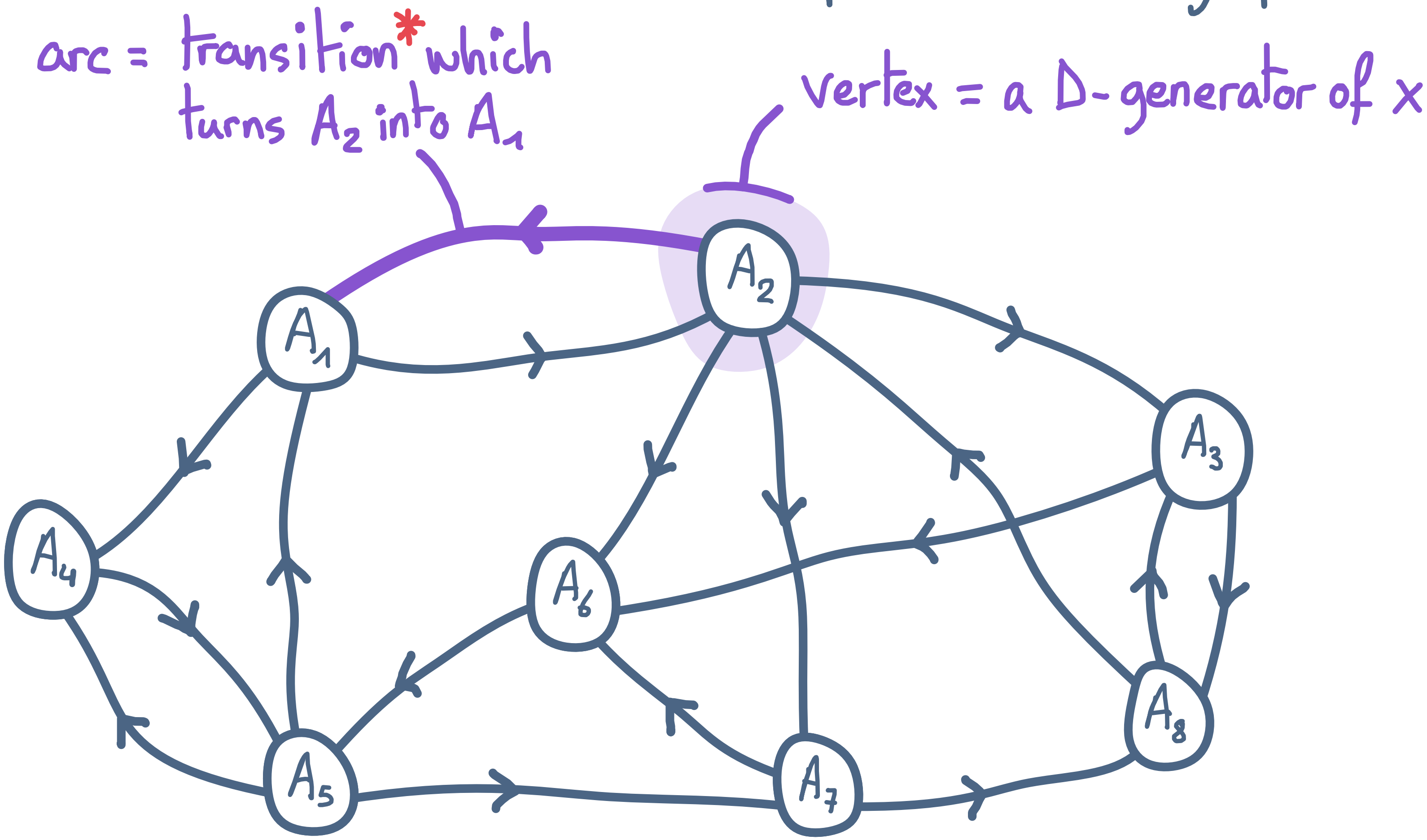
- output quasi-poly from $\mathcal{M}_i$
- poly-delay from Σ

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

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

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

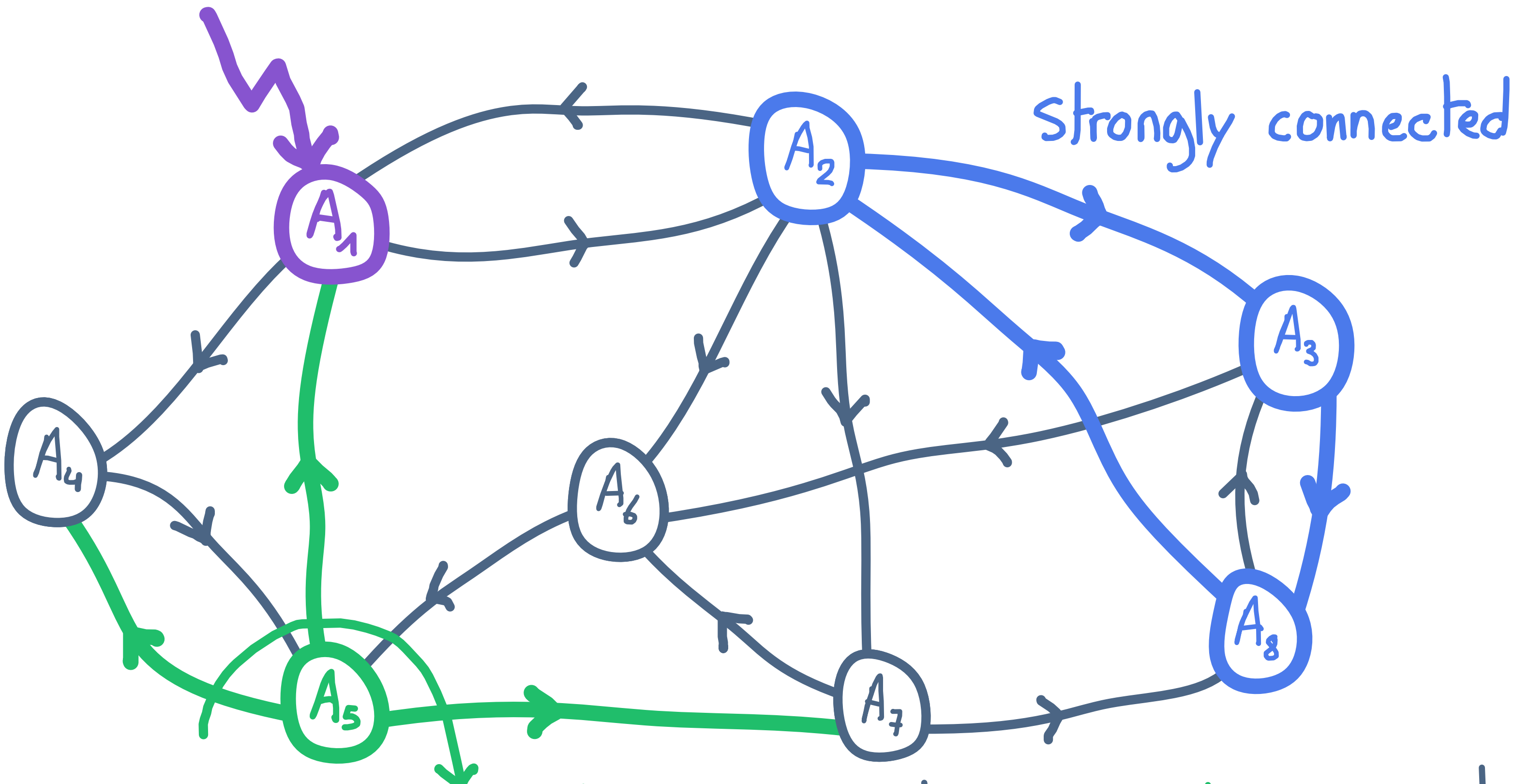
Principle: Solution-graph traversal



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

(greedily) minimize w.r.t. Σb

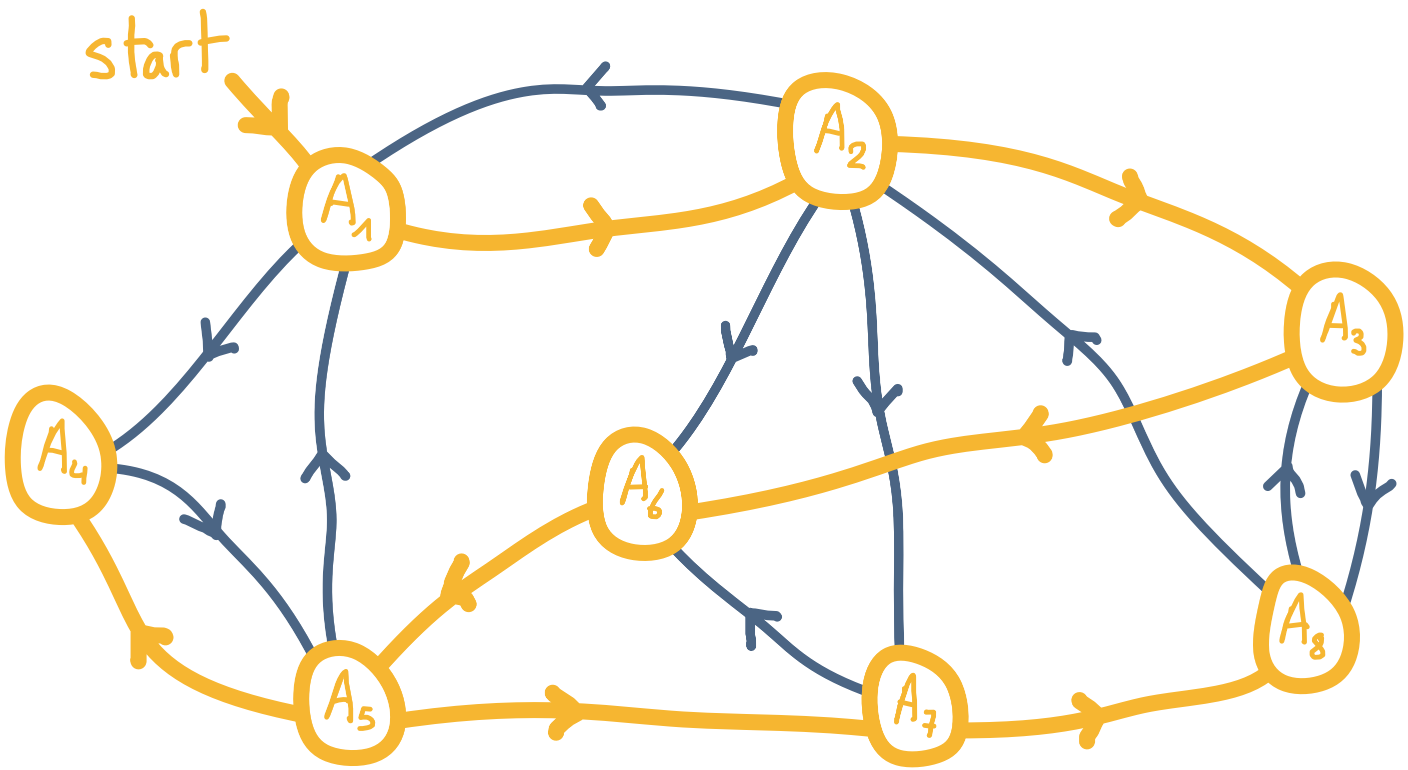
Principle: Solution-graph traversal



poly number of transitions poly-time computable

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

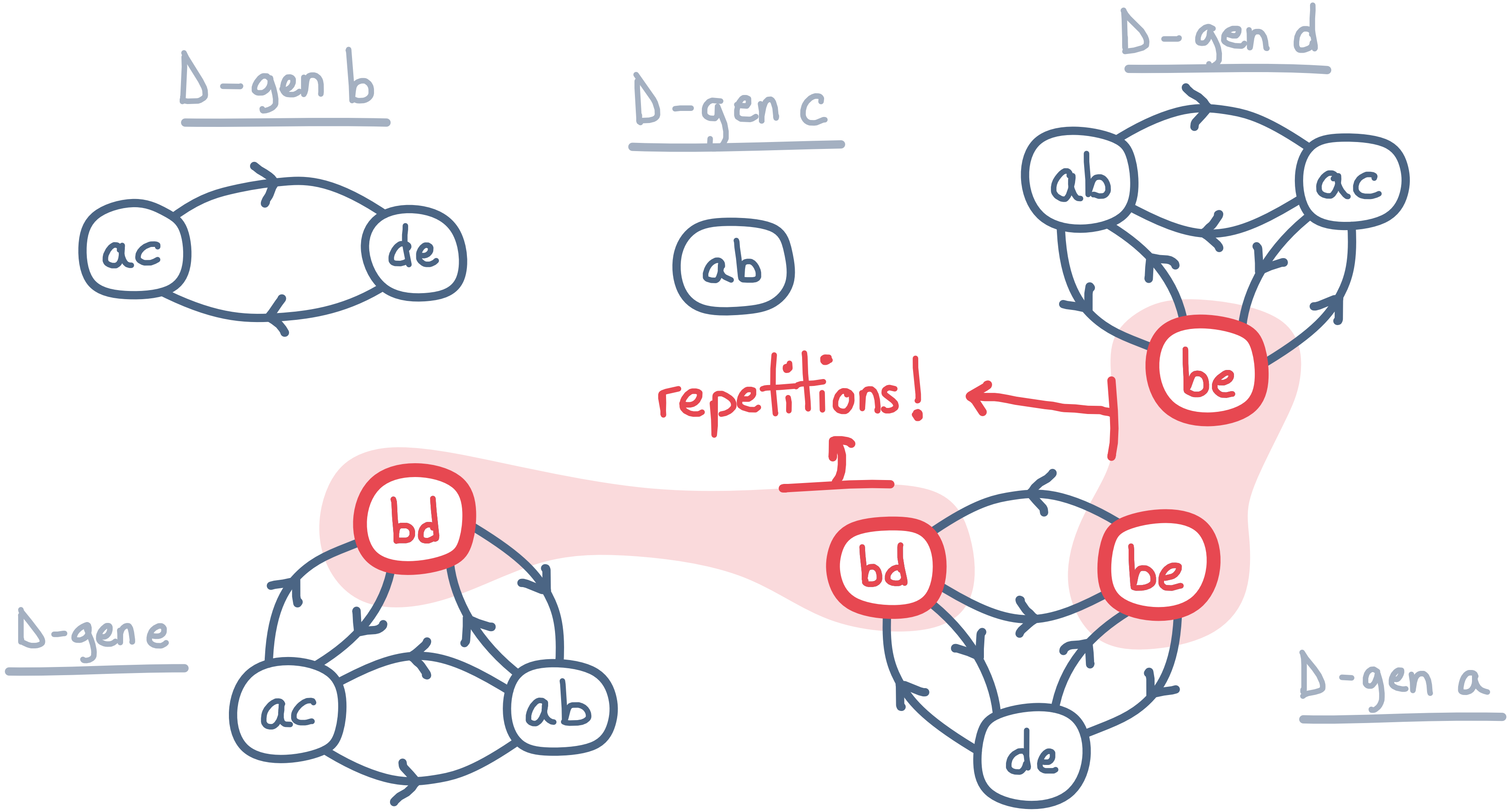
Principle: Solution-graph traversal



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

$\Rightarrow$ poly-delay enumeration (with DFS) of D-gen of some x

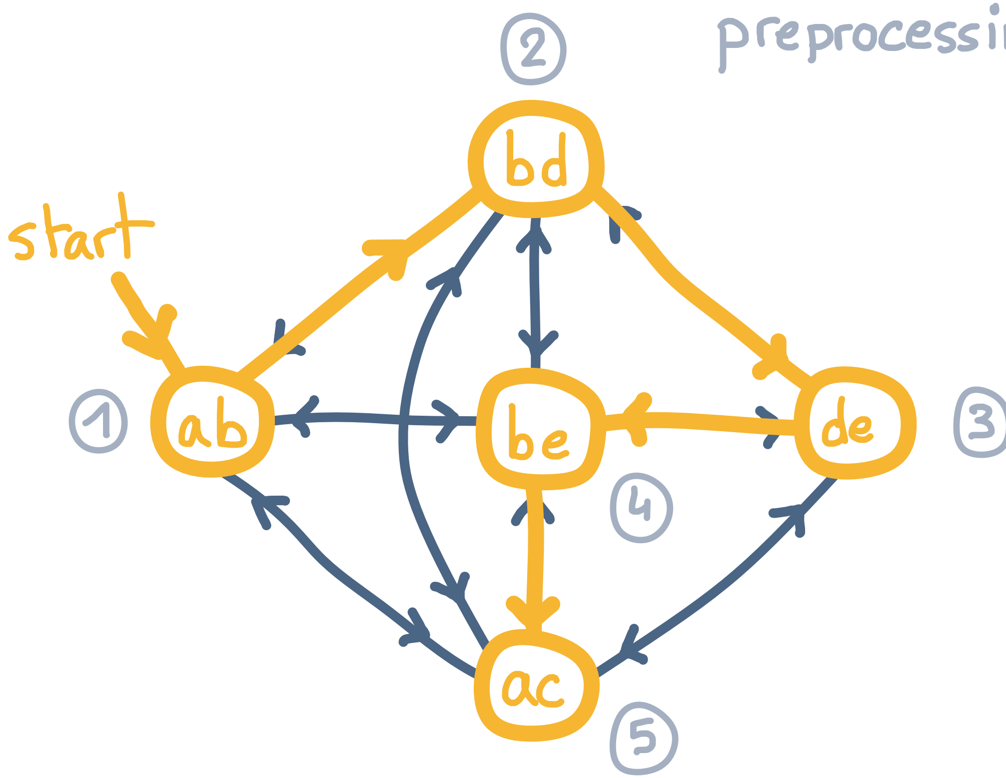
In our case (running ex)



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

$\Rightarrow$ no guarantee on delay

Fix: merge the graphs



preprocessing

- ⓪ $d \to c, \; e \to c$
- ① $ab \to c, \; ab \to d, \; ab \to e$
- ② $bd \to a, \; bd \to e$
- ③ $de \to a, \; de \to b$
- ④ $be \to a, \; be \to d$
- ⑤ $ac \to b, \; ac \to d, \; ac \to e$

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+

References

Dilworth

Lattices with unique irreducible decompositions

Annals of Mathematics, 1940

Finkbeiner

A general dependence relation for lattices

Proceedings of the AMS, 1951

Jónsson, Nation

A report on sublattices of a free lattice

Contributions to Universal Algebra, 1977

Dilworth, 40

Finkbeiner, 51

Jónsson, Nation, 77

References

Gaskill, Rival, 78

Gaskill, Rival

An exchange property for modular lattices

Algebra Universalis, 1978

Day, 79

Day

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

Canadian Journal of Mathematics

Faigle, 86

Faigle

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, 22

Elbassioni
On dualization over distributive lattices
Discrete Mathematics and Theoretical Computer Science, 2022

Kavvadias et al., 00

Kavvadias, Sideri, Stavropoulos
Generating maximal models of a Boolean expression
Information Processing Letters, 2000