Towards concept analysis in categories: limit inferior as algebra, limit superior as coalgebra

Toshiki Kataoka
The University of Tokyo,
JSPS Research Fellow

&

Dusko Pavlovic
University of Hawaii at Manoa

Overview

- Concept analysis
  - Concept analysis *in categories* — new problem
- Dedekind–MacNeille completion
  - Generalizations of Dedekind–MacNeille completion
  - Bicompletions *of categories* — new answer

Concept Analysis

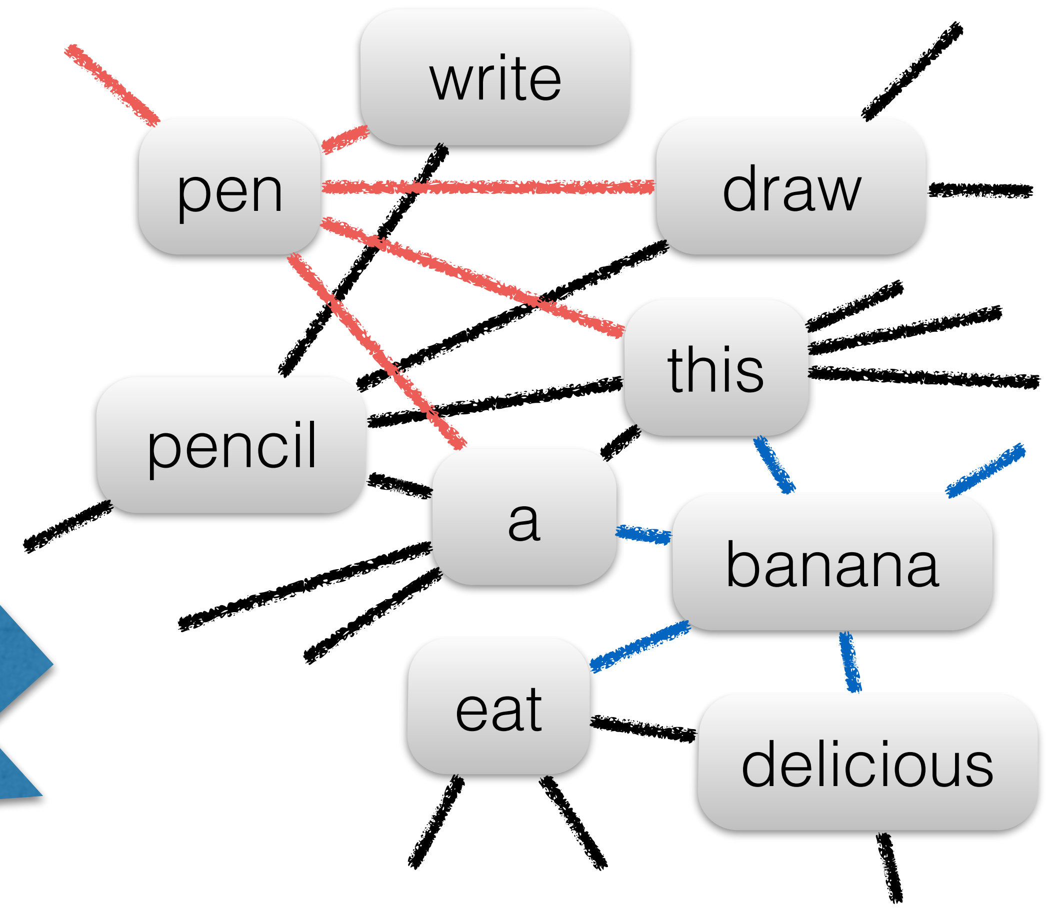
(or, knowledge acquisition,
semantic indexing,
data mining)

Example: Text Analysis

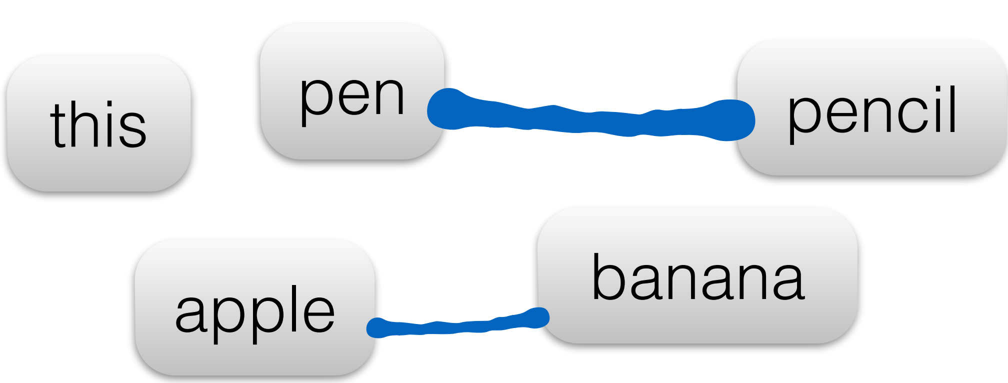
Corpus

This is a pen. Is that a pen?
…

Co-occurrence



Proximity among words



Concept Analysis

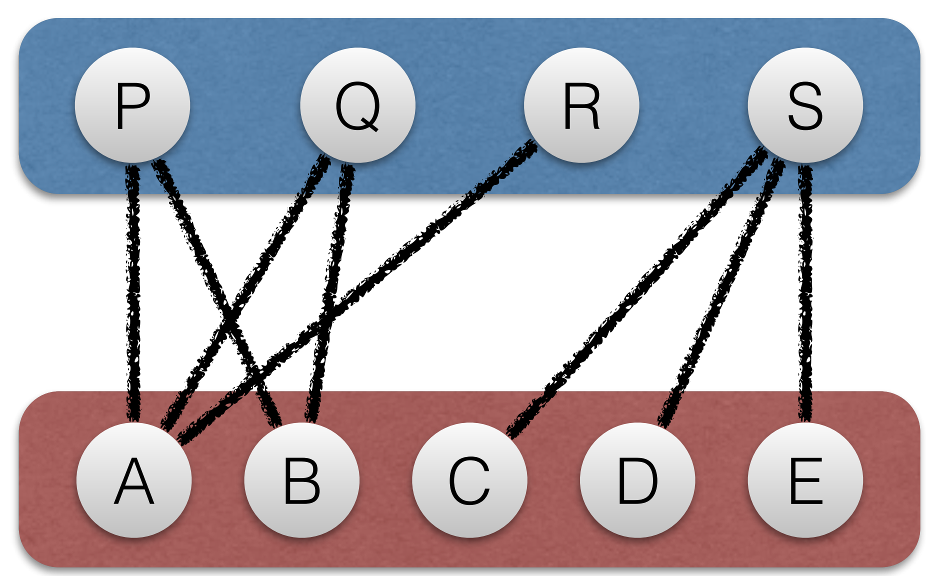
- Given data in a **matrix**

between two

- Extract information as **vectors**

either side

	P	Q	R	S
Alice	✔	✔	✔	
Bob	✔	✔		
Carol				✔
Dan				✔
Eve				✔



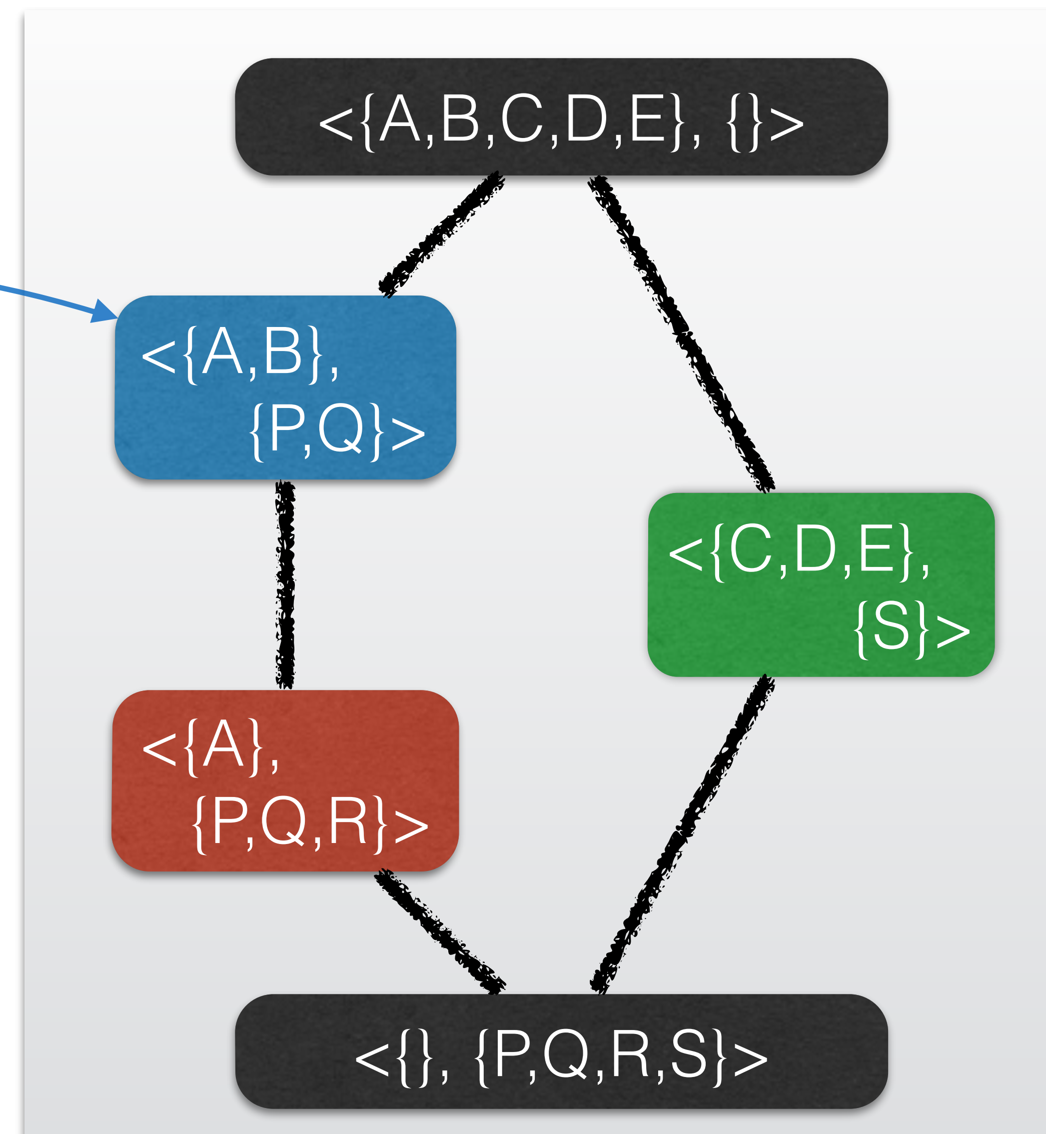
Formal Concept Analysis

[Wille, 1982]

- Preordered

	P	Q	R	S
Alice	✔	✔	✔	
Bob	✔	✔		
Carol				✔
Dan				✔
Eve				✔

Concept lattice



Latent Semantic Analysis
(Principal Component Analysis)

- Linear algebraic

	P	Q	R	S
Alice	★★	★★★	★★★★	
Bob	★★	★★★★★		
Carol				★★★★★
Dan				★★★★
Eve				★

Column weights: $[\,1 : 2 : 1 : 0\,]$; row weights: $[\,0 : 0 : 0 : 1 : 1\,]$

$$M = \frac{1}{5}\begin{pmatrix} 2 & 3 & 4 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{\dagger}$$

$$M = U\Sigma V^{\dagger}$$

where

$$U = \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \frac{\sqrt{48}}{5} & 0 & 0 \\ 0 & \frac{\sqrt{42}}{5} & 0 \\ 0 & 0 & \frac{\sqrt{10}}{5} \end{pmatrix}$$

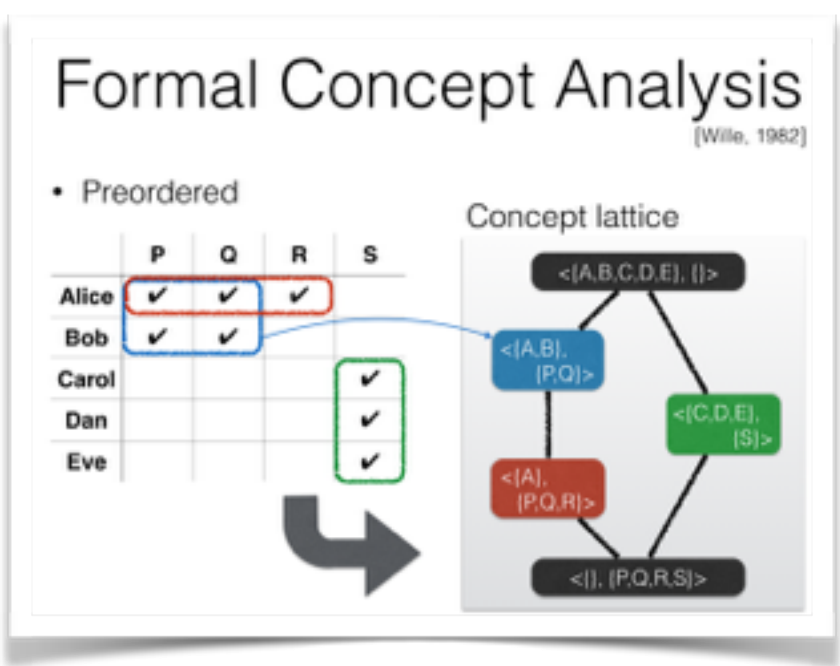
$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{42}} & 0 \\ 0 & \frac{4}{\sqrt{42}} & 0 \\ 0 & \frac{1}{\sqrt{42}} & 0 \end{pmatrix}$$

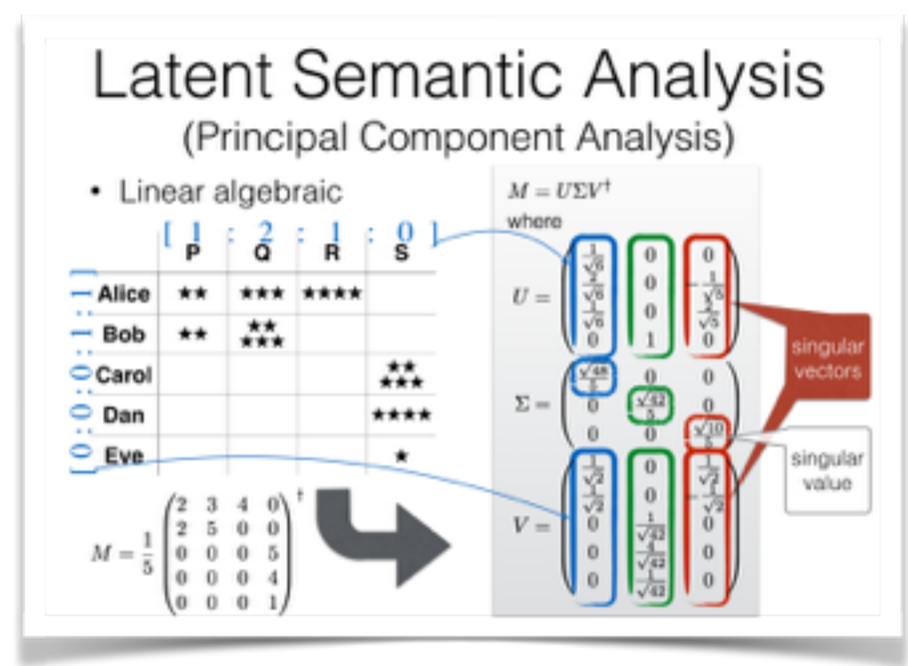
singular vectors

singular value

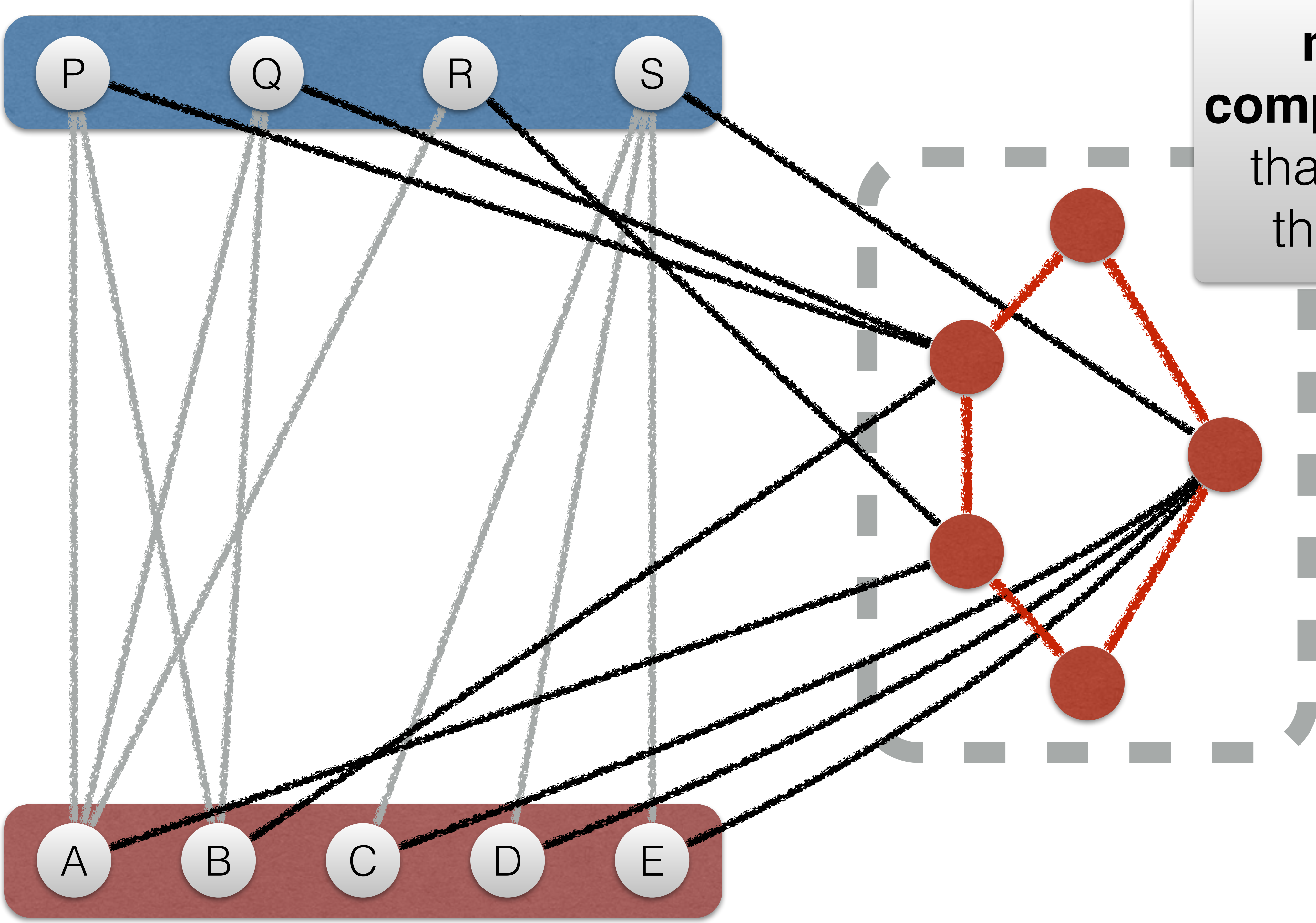
Towards Unification

- Fixed points (definition)
- Completeness (theorem)
- Minimality (theorem)





Factorizations of Relations

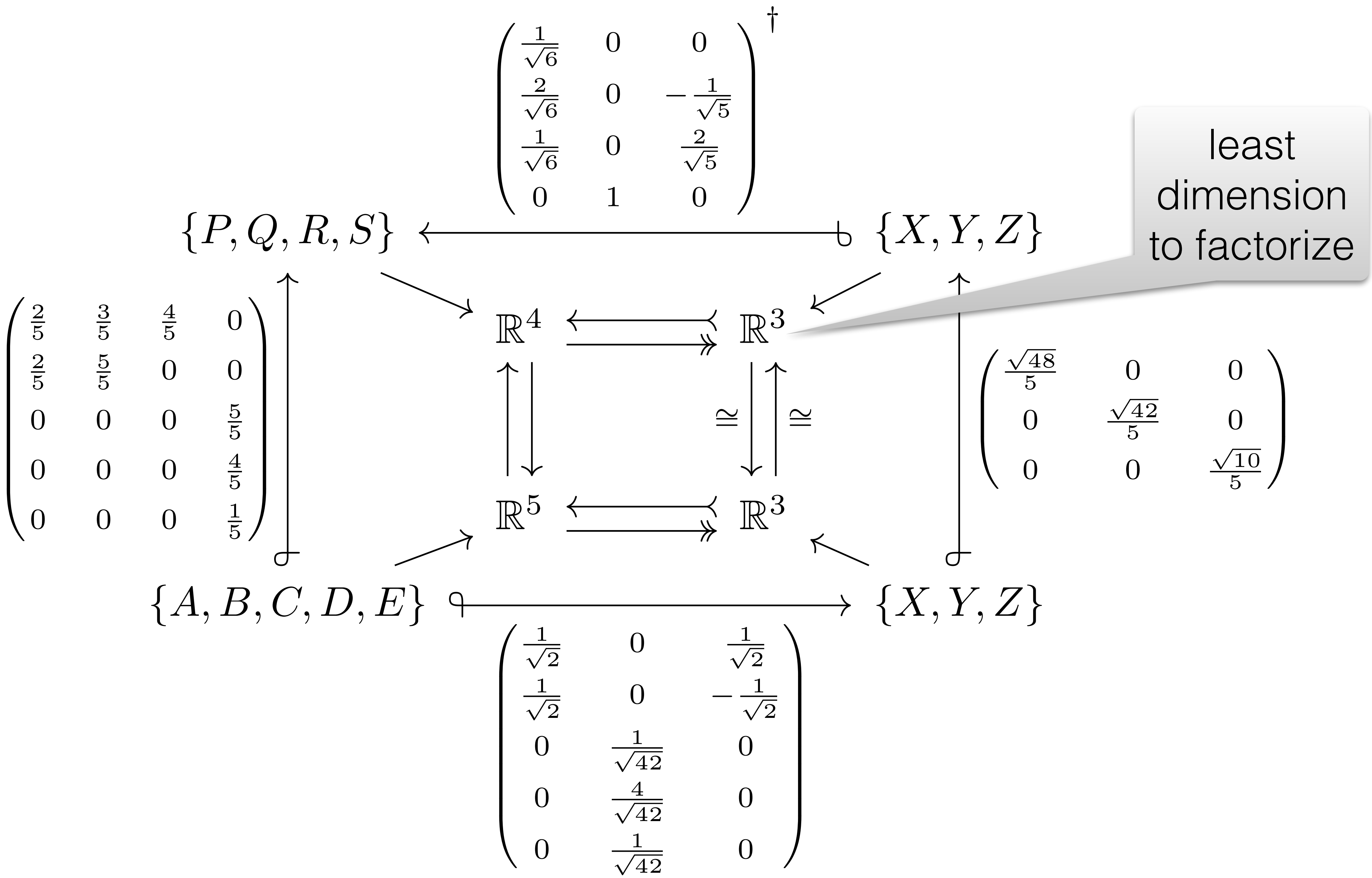


minimal complete lattice that factorize the relation

Singular Value Decomposition

- Compact SVD: $M = U\Sigma V^\dagger$
  - M: (given) real $m\times n$ matrix, rank r
  - U: $m\times r$ matrix, $U^\dagger U = I_r$
  - V: $n\times r$ matrix, $V^\dagger V = I_r$
  - Σ: diagonal $r\times r$ matrix
    - with positive reals on the diagonal

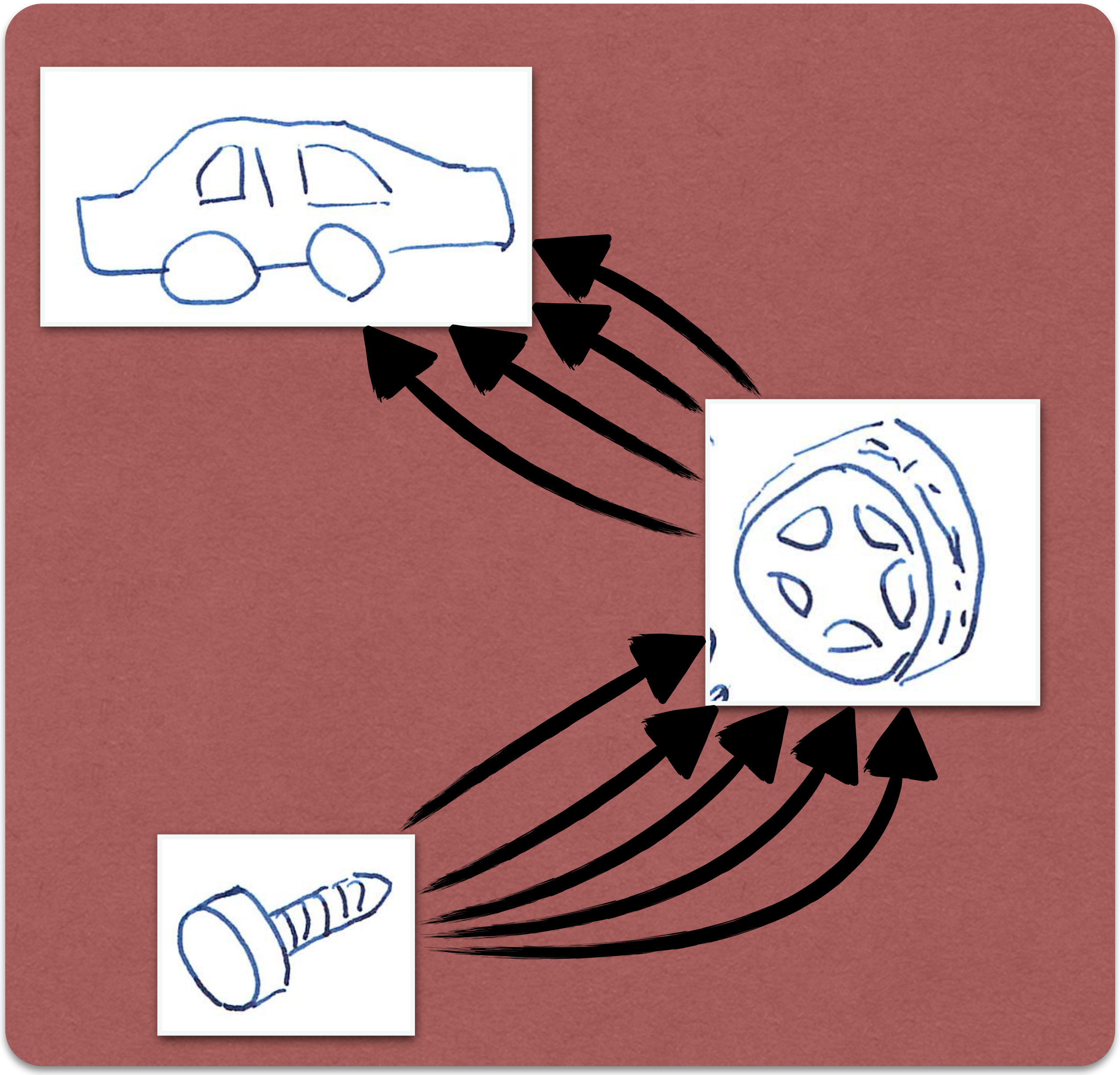
Factorizations of Real Matrices



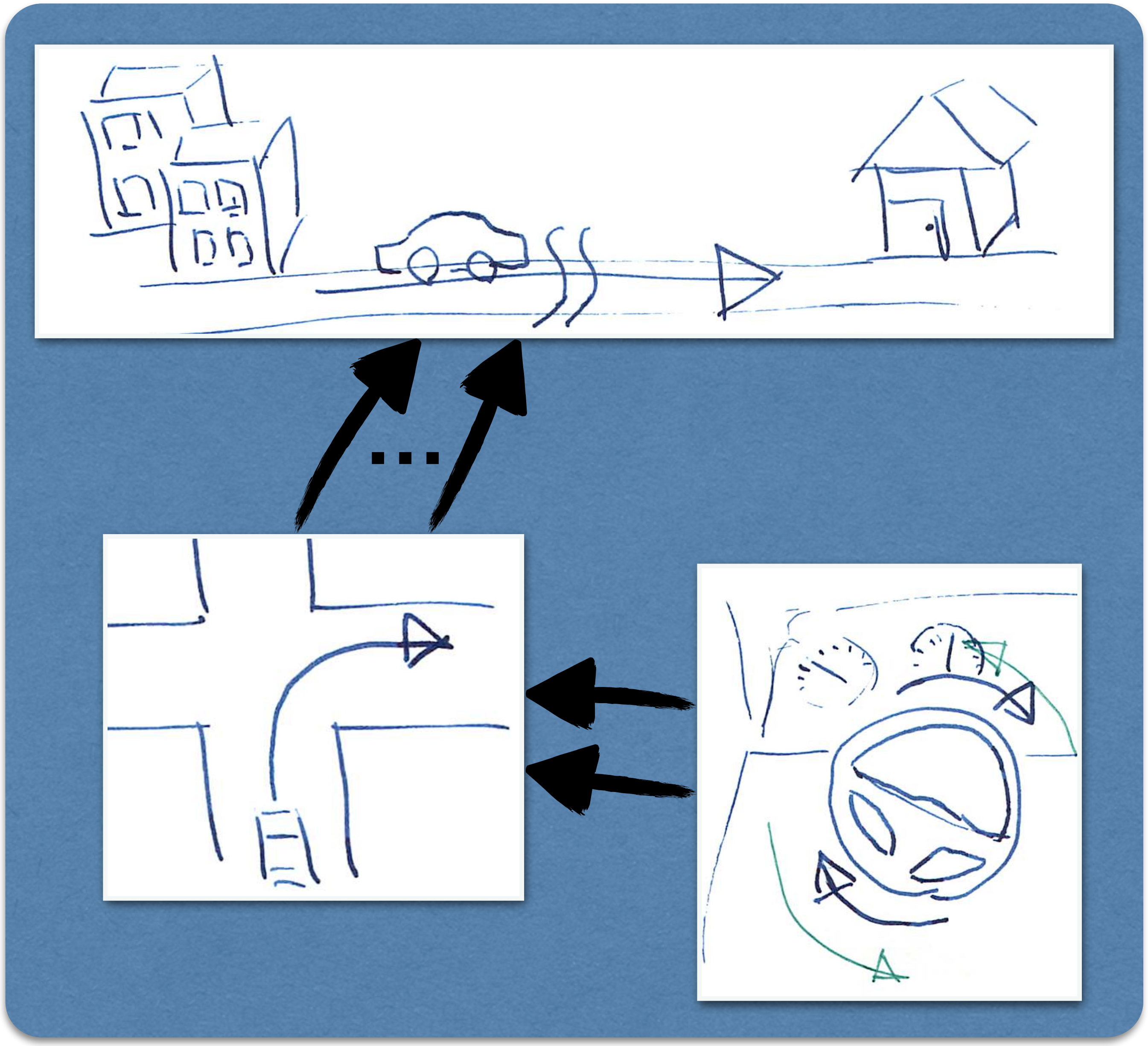
Concept Analysis
in Categories

Components and Functionalities

Structural components



Functional modules



Engineering

- Functionalities: top-down
- Components: bottom-up
- Relation: specification

V-model

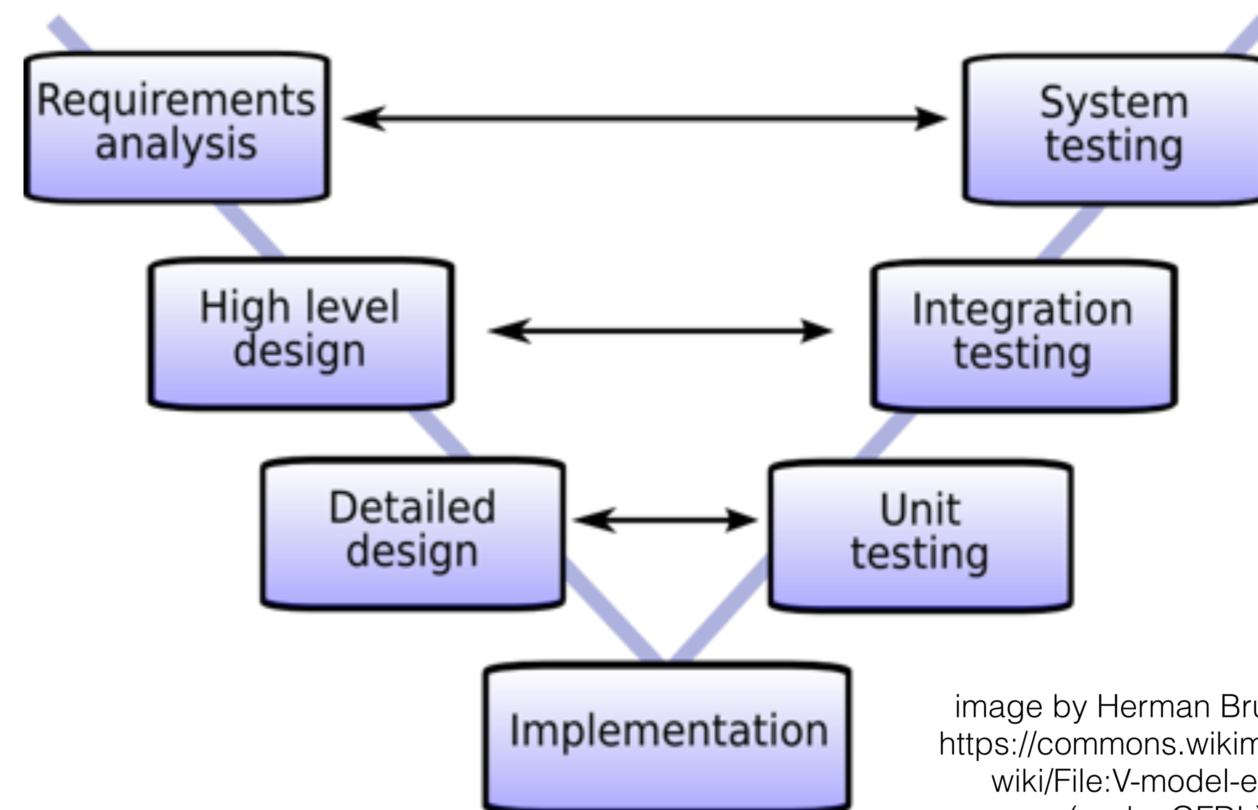
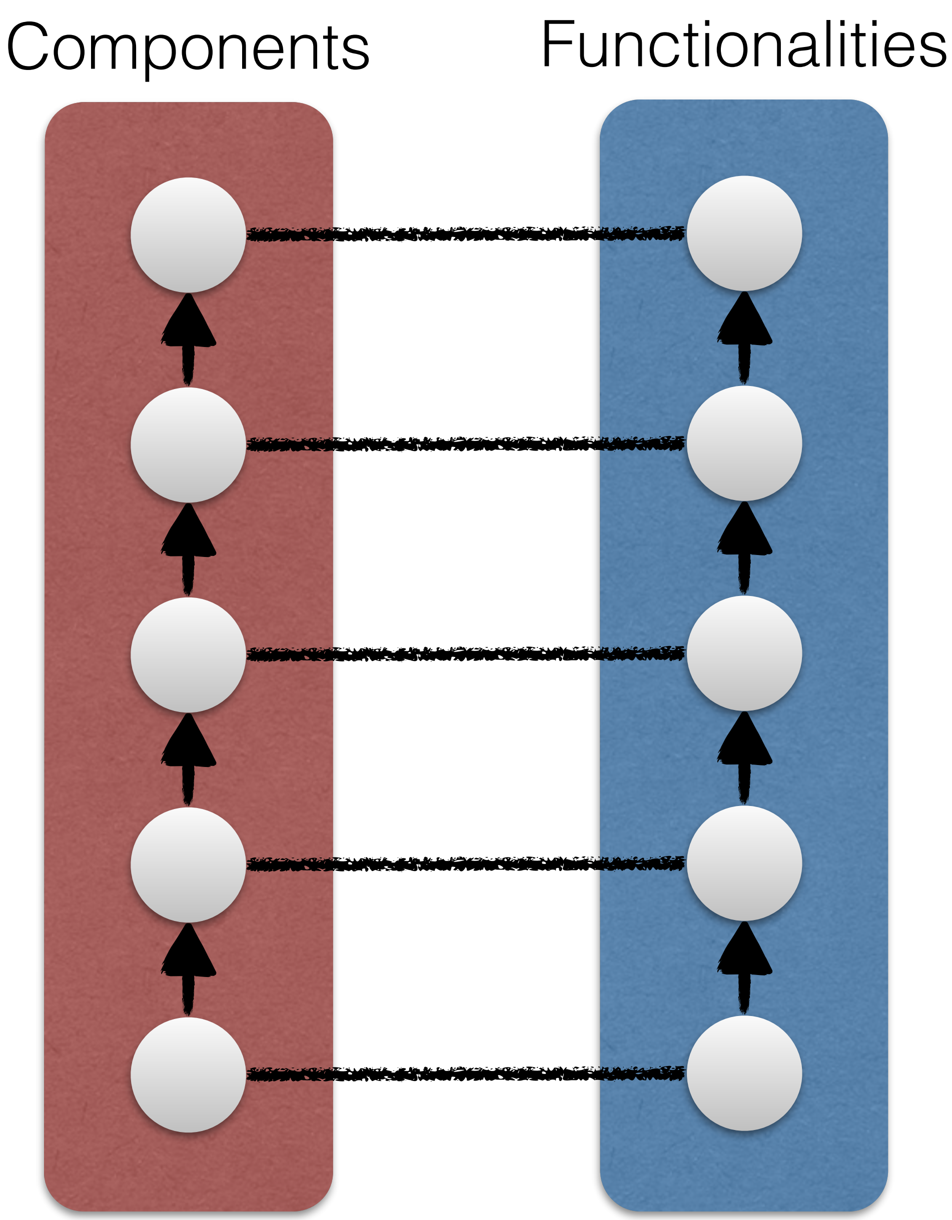


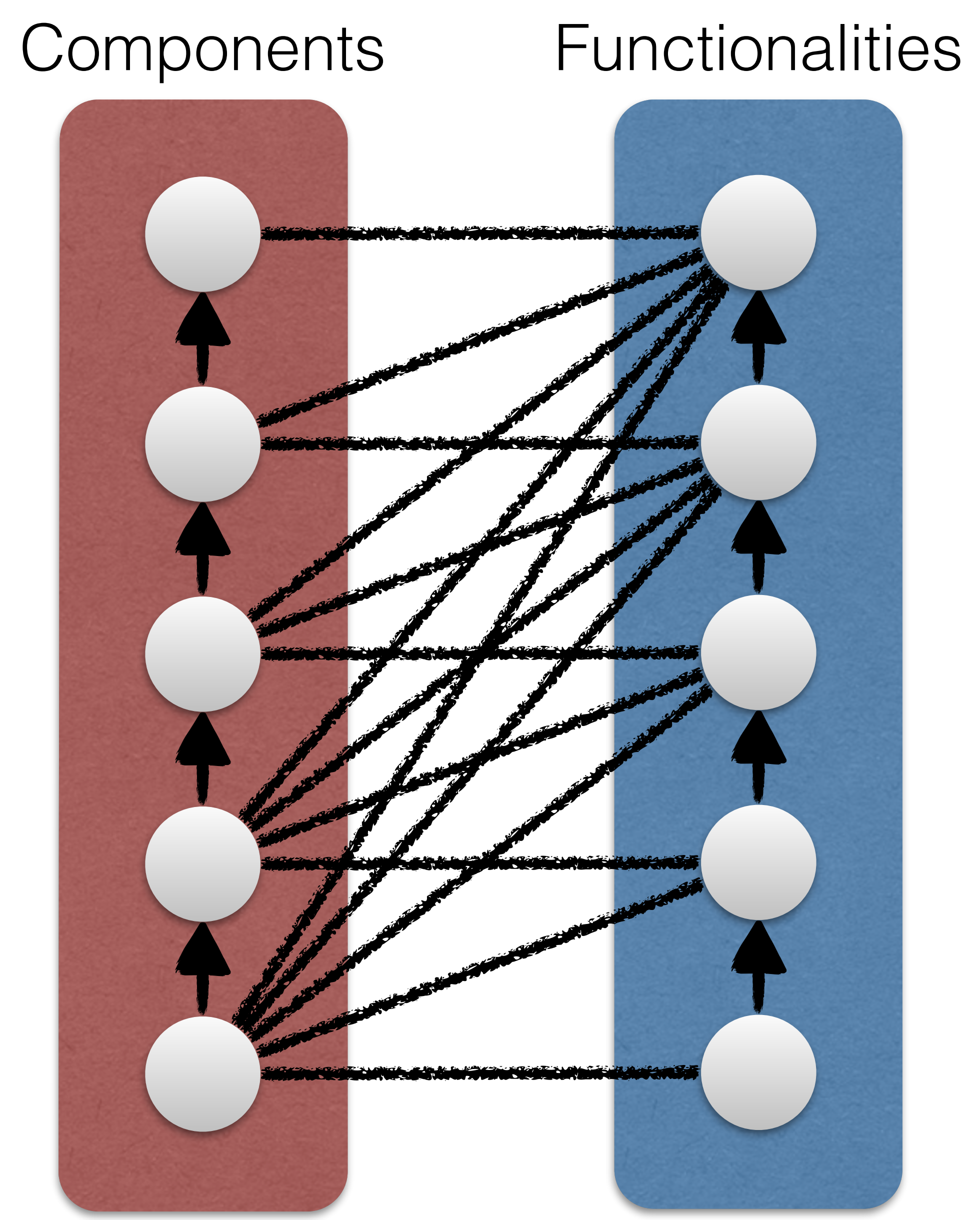
image by Herman Bruyninckx
https://commons.wikimedia.org/wiki/File:V-model-en.png
(under GFDL)

Reverse-Engineering

- Want to know

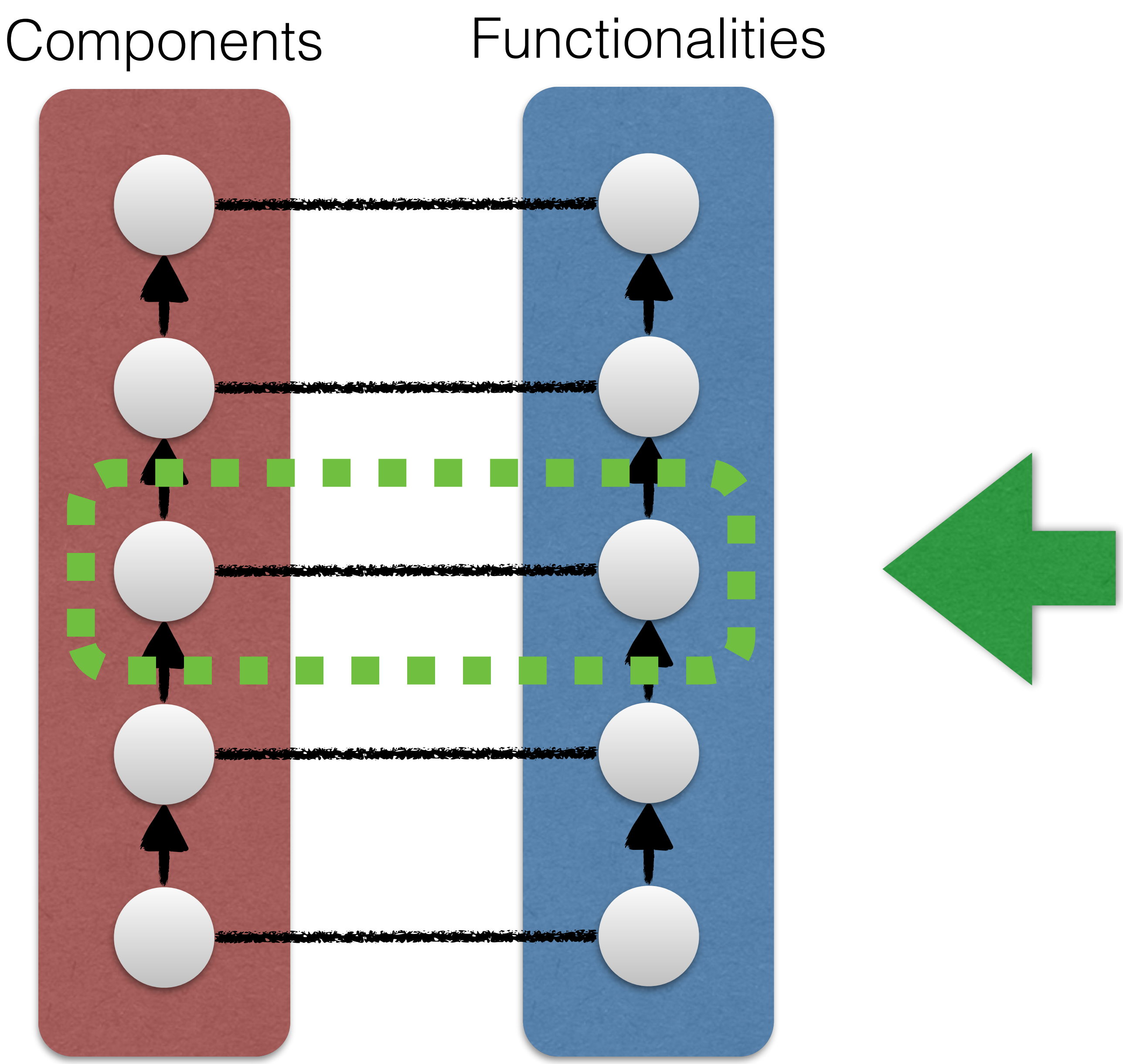


- Can be observed

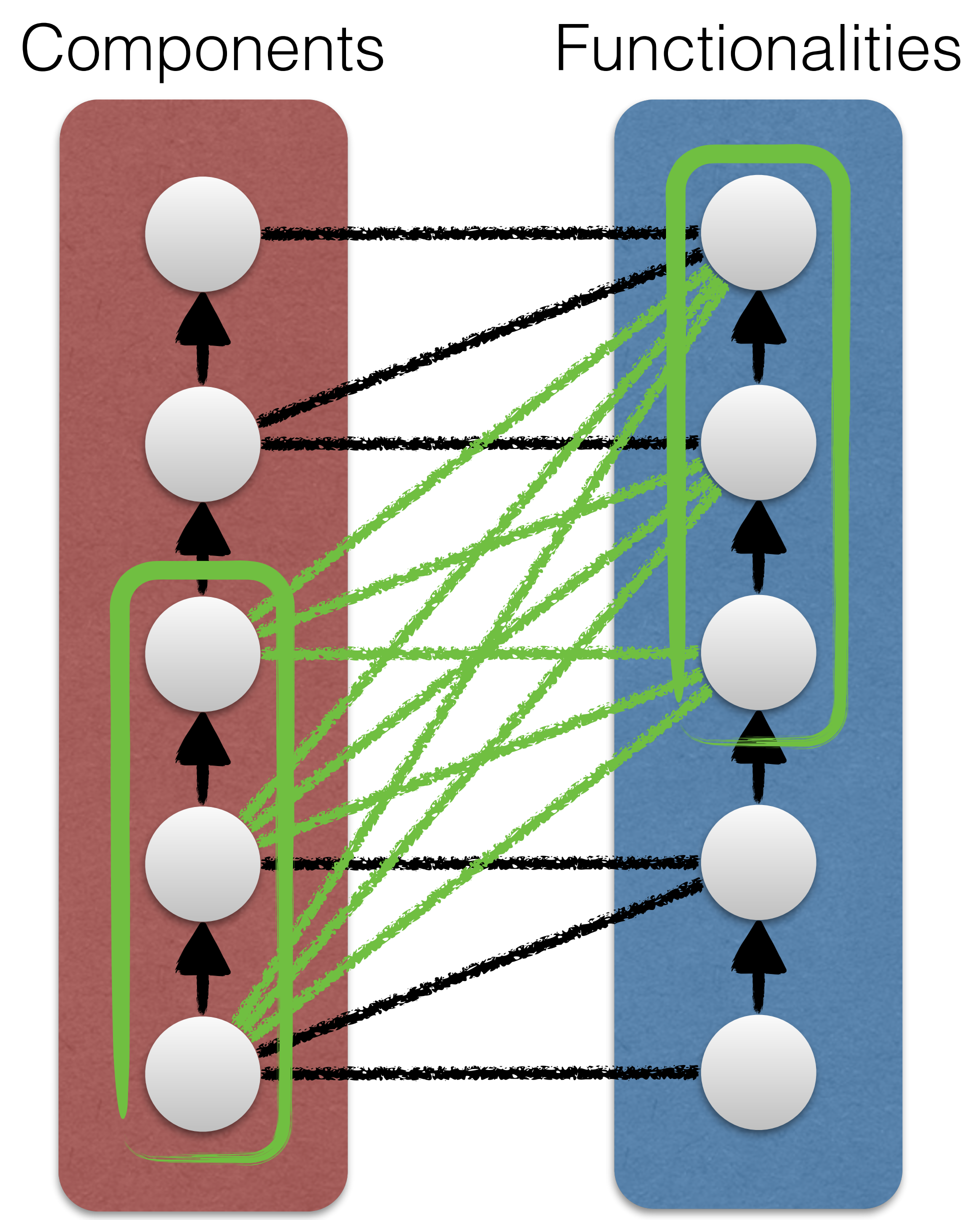


Reverse-Engineering

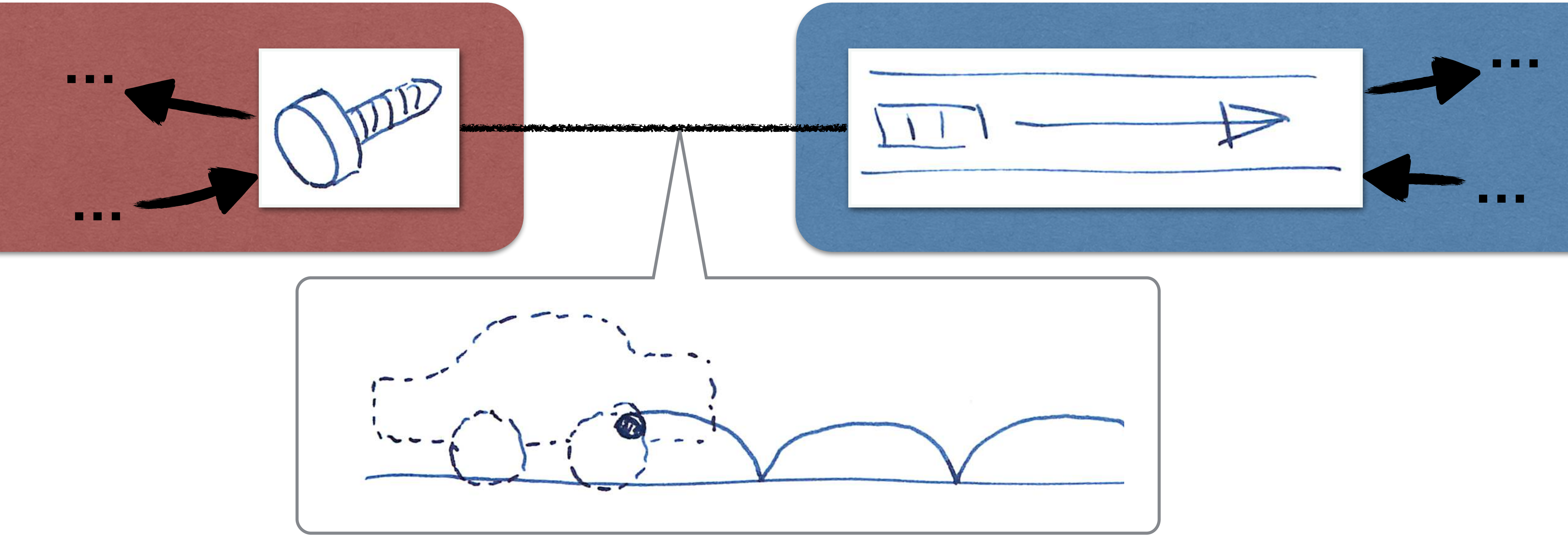
- Want to know



- Can be observed



Actions on Relations



- From super-components
- From sub-functionalities

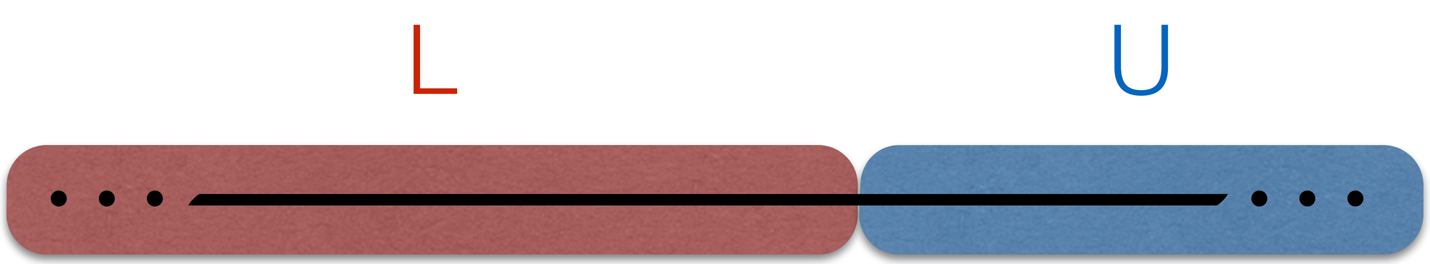
$\Phi\colon \mathbb{A}^{\mathrm{op}} \times \mathbb{B} \to \mathbf{Set}$

Dedekind–MacNeille Completion

Dedekind Cuts [Dedekind, 1872]

$$\mathbb{R} = \{\langle L, U\rangle \mid \mathbb{Q} = L \amalg U,\ \ L, U \neq \emptyset,\ \forall l \in L.\, \forall u \in U.\, l \leq u\}/\sim$$

where $\langle\{< q\}, \{\geq q\}\rangle \sim \langle\{\leq q\}, \{> q\}\rangle$ $(q \in \mathbb{Q})$



Dedekind–MacNeille Completion [MacNeille, 1937]

- $(P, \leq)$: poset

$$\updownarrow P = \{\langle L, U\rangle \mid L \subseteq P,\ U \subseteq P,\ L = \{l \in P \mid \forall u \in U.\ l \leq u\},\ U = \{u \in P \mid \forall l \in L.\ l \leq u\}\}$$

lower bounds

upper bounds

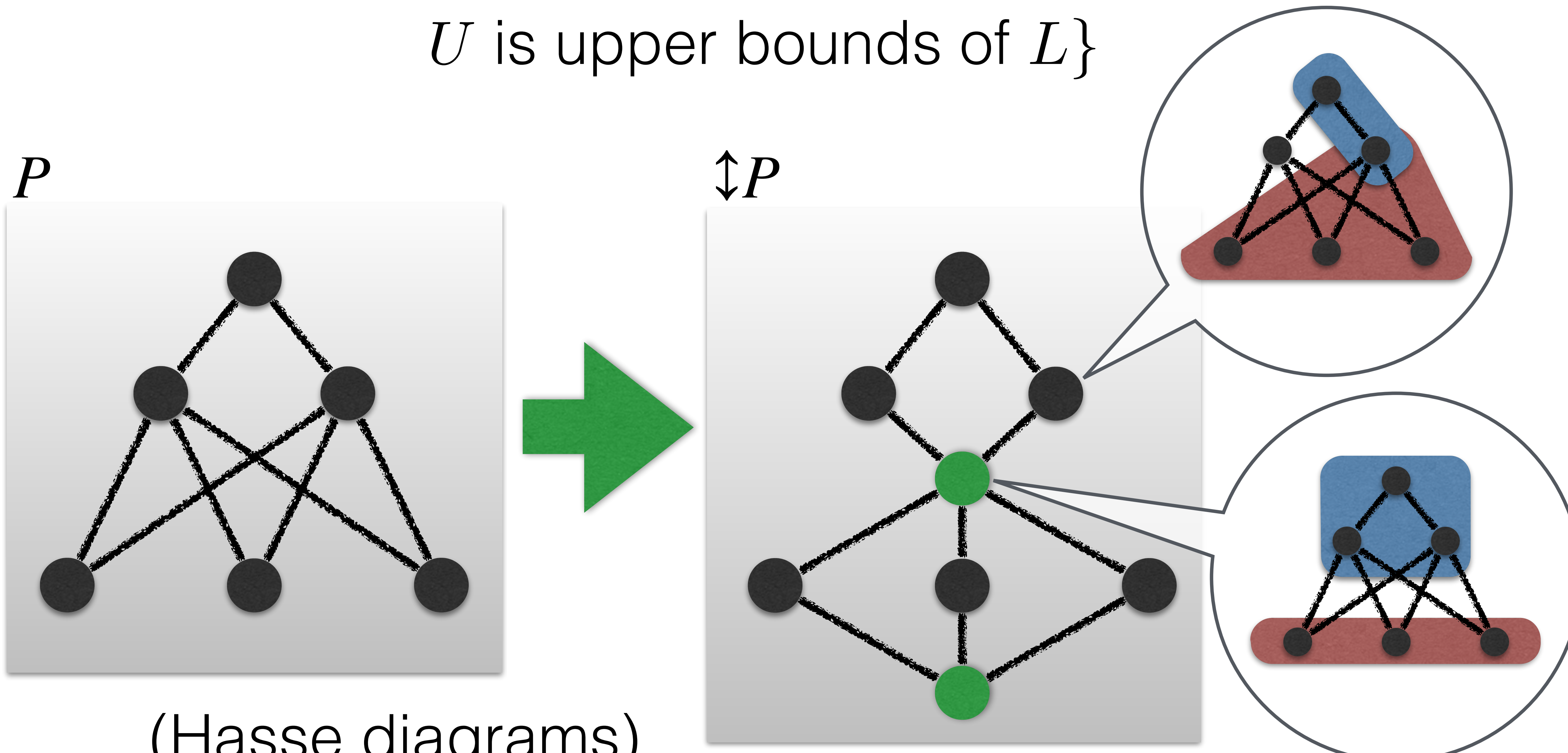
$$\langle L, U\rangle \leq \langle L', U'\rangle \iff L \subseteq L' \iff U \supseteq U'$$

- Generalizes Dedekind cuts

$$\updownarrow \mathbb{Q} = \{\langle \emptyset, \mathbb{Q}\rangle\} \cup \{\langle \{\leq r\} \cap \mathbb{Q}, \{\geq r\} \cap \mathbb{Q}\rangle \mid r \in \mathbb{R}\} \cup \{\langle \mathbb{Q}, \emptyset\rangle\}$$
$$\cong \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$$

An Example

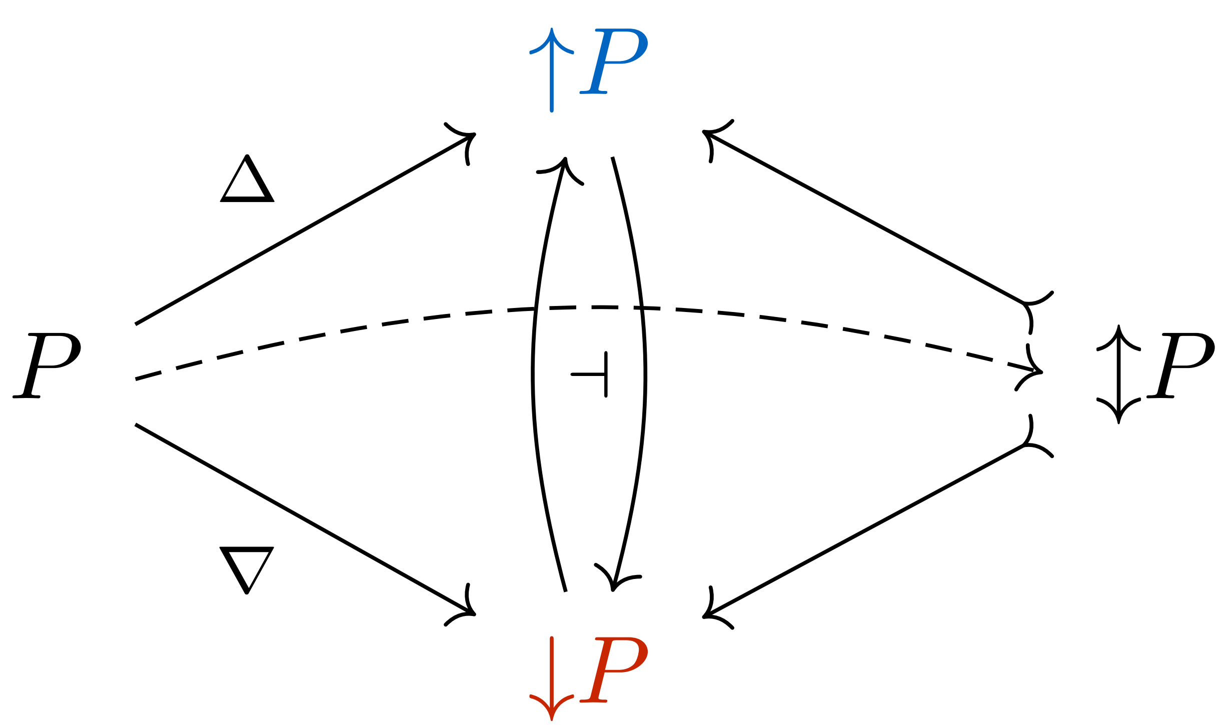
$$\updownarrow P = \{\langle L, U\rangle \mid L \text{ is lower bounds of } U,\ U \text{ is upper bounds of } L\}$$



(Hasse diagrams)

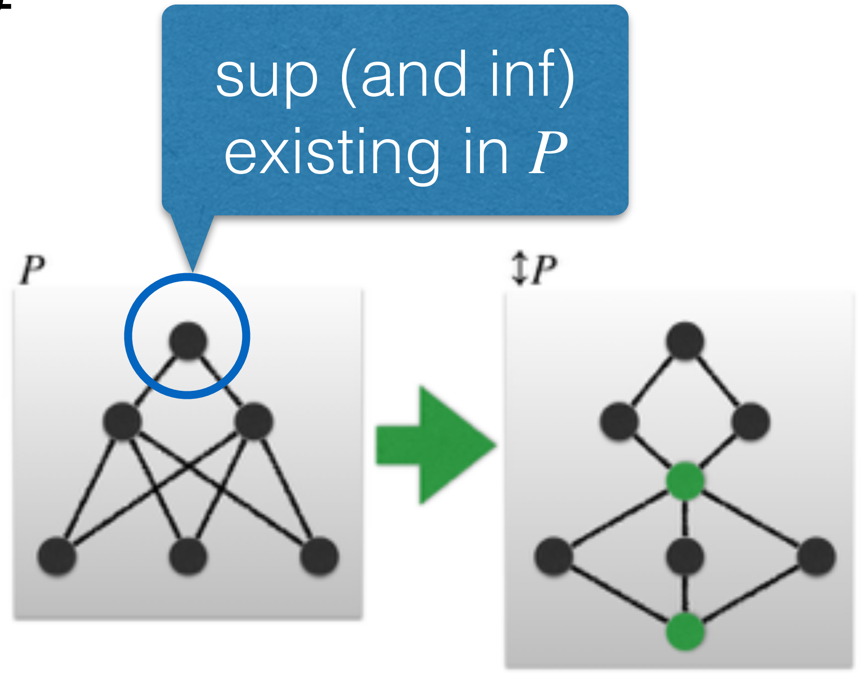
Categorically

- $(P, \leq)$: poset,
- $(\downarrow P, \subseteq)$ (resp. $(\uparrow P, \supseteq)$): the family of lower (resp. upper) sets
- $\updownarrow P$: fixed point of the Galois connection



Properties

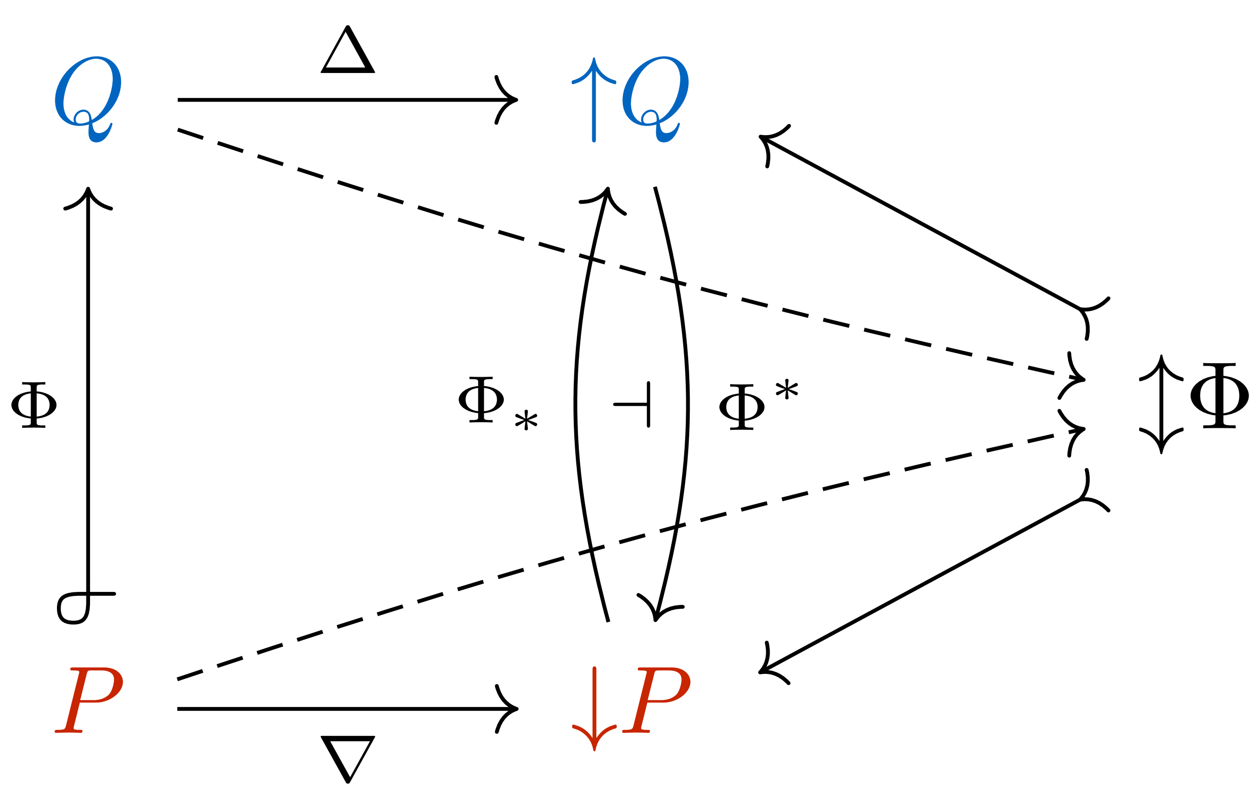
- $\updownarrow P$ is inf- and sup-complete. (Complete lattice)
- $P \to \updownarrow P$ is an inf- and sup-dense embedding.
  - Thus, sup- and inf-preserving
- The **minimal bicompletion**

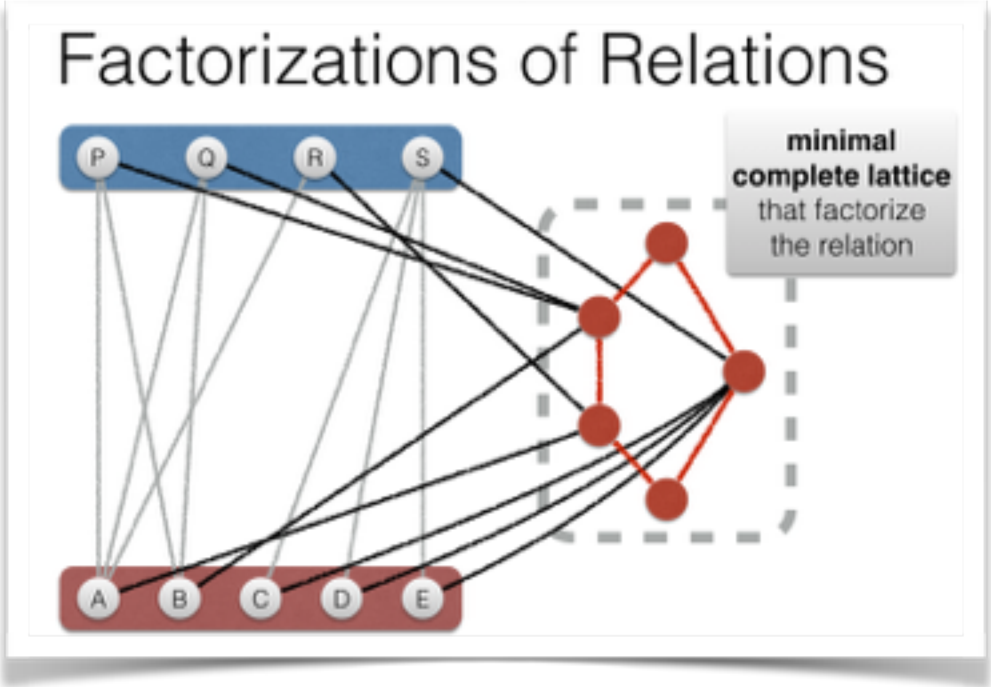


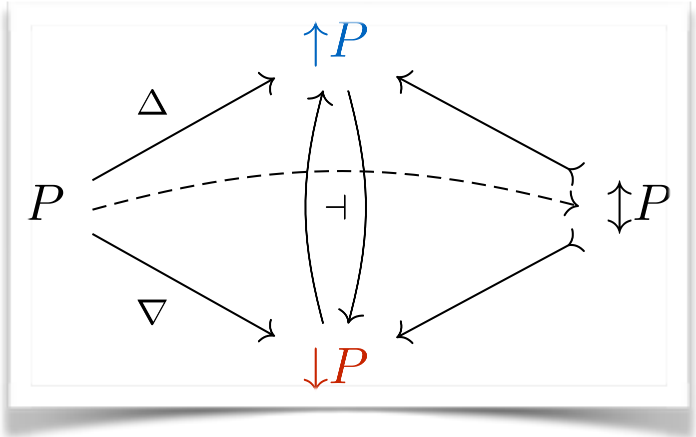
Generalizations of Dedekind–MacNeille Completion

Completions of Relations

- $\Phi \subseteq P \times Q$: relation s.t. $x' \leq x,\ x\Phi y,\ y \leq y' \Rightarrow x'\Phi y'$





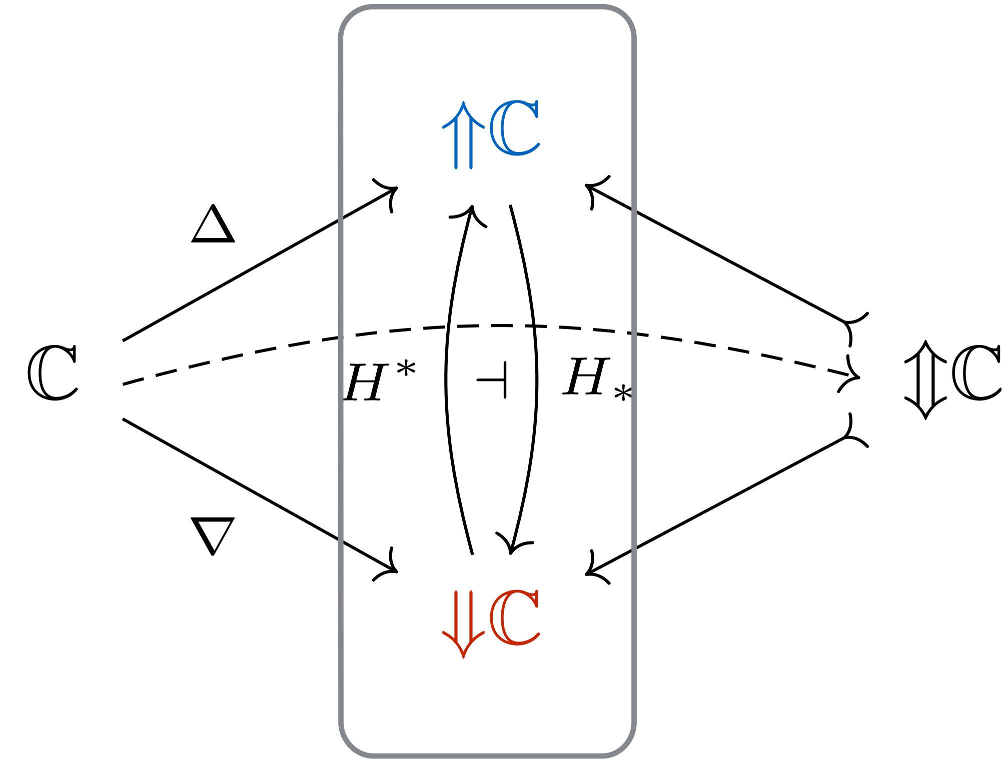


Notations

- $\mathbb{C}$: $\mathcal{V}$-enriched category,
- $\Downarrow\mathbb{C} = [\mathbb{C}^{\text{op}}, \mathcal{V}]$: category of presheaves,
- $\Uparrow\mathbb{C} = [\mathbb{C}, \mathcal{V}]^{\text{op}}$: category of postsheaves,
- $\nabla\colon \mathbb{C} \to \Downarrow\mathbb{C}$, $\Delta\colon \mathbb{C} \to \Uparrow\mathbb{C}$: Yoneda embeddings,
- $\mathbb{A} \looparrowright \mathbb{B}$ denotes $\mathbb{A}^{\text{op}} \otimes \mathbb{B} \to \mathcal{V}$.

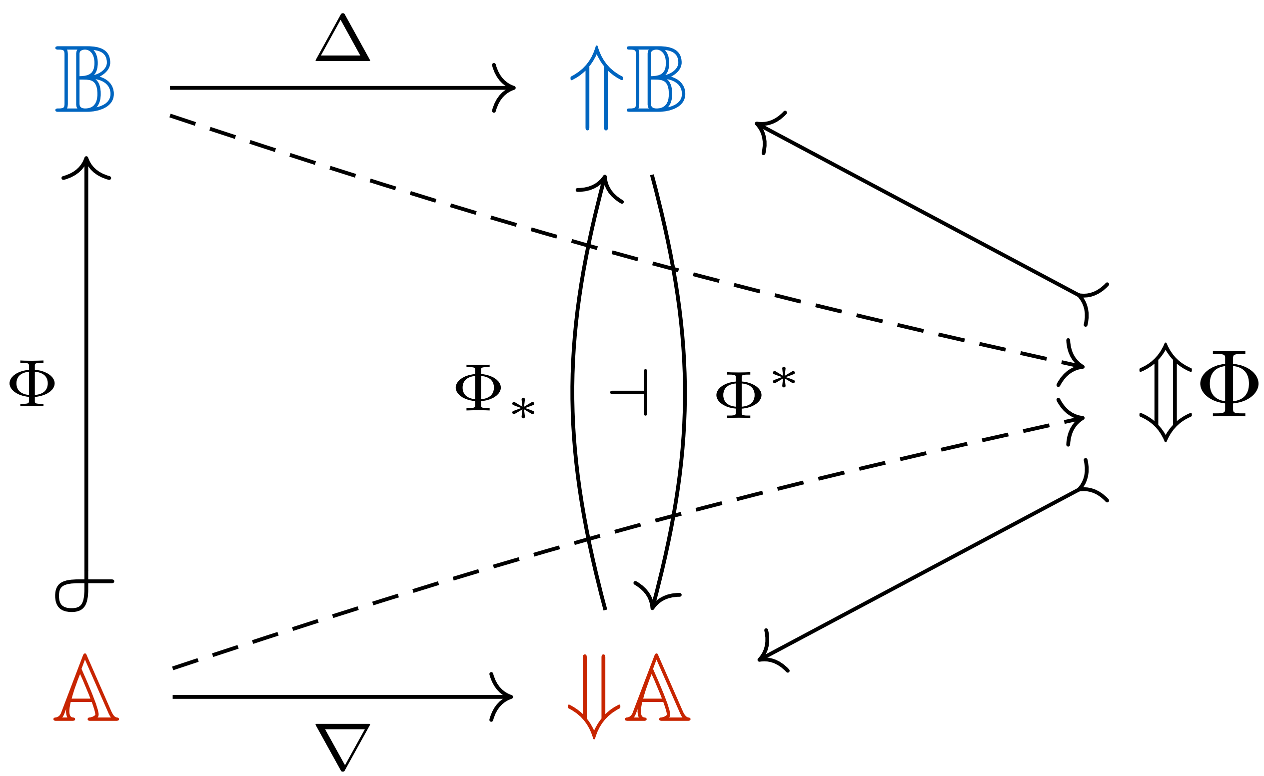
Isbell Completion

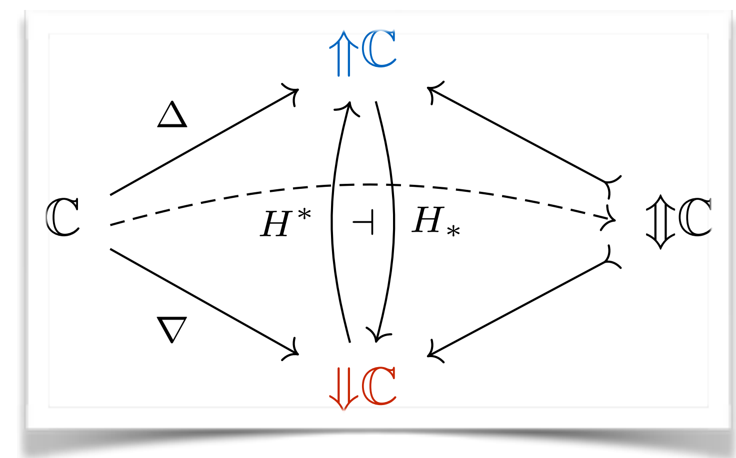
- $\mathcal{V} = \{0,1\}$ ($\mathcal{V}$-category = poset)
    - $\Updownarrow\mathbf{C}$: Dedekind–MacNeille completion
- $\mathcal{V} = [0,\infty]$ ($\mathcal{V}$-category = Lawvere metric space)
    - $\Updownarrow\mathbf{C}$: (directed) tight span [Willerton, 2013]

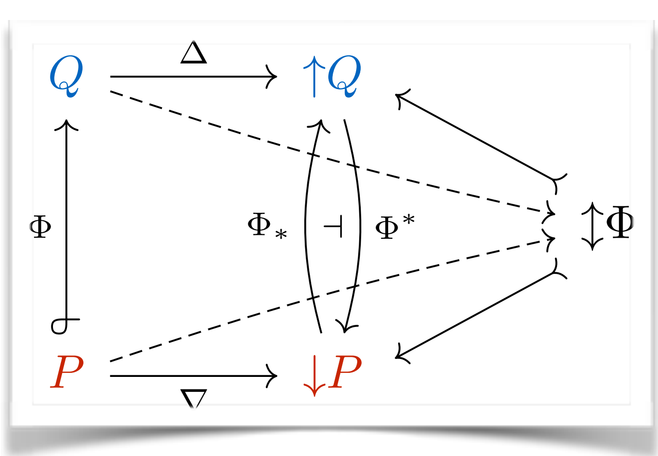


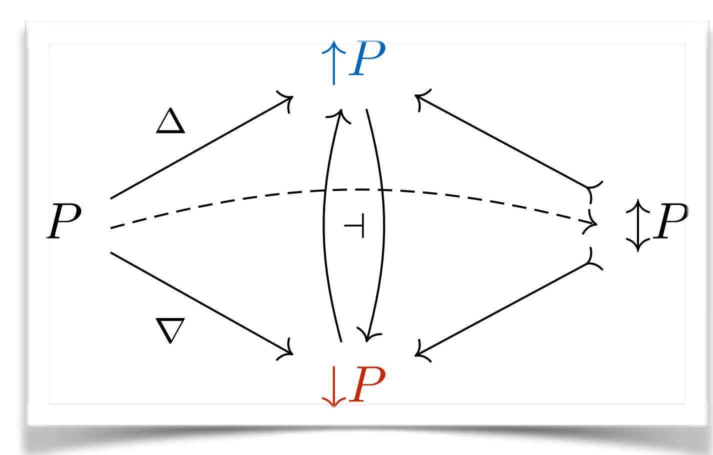
Isbell duality

If $\mathcal{V} = [0,\infty]$









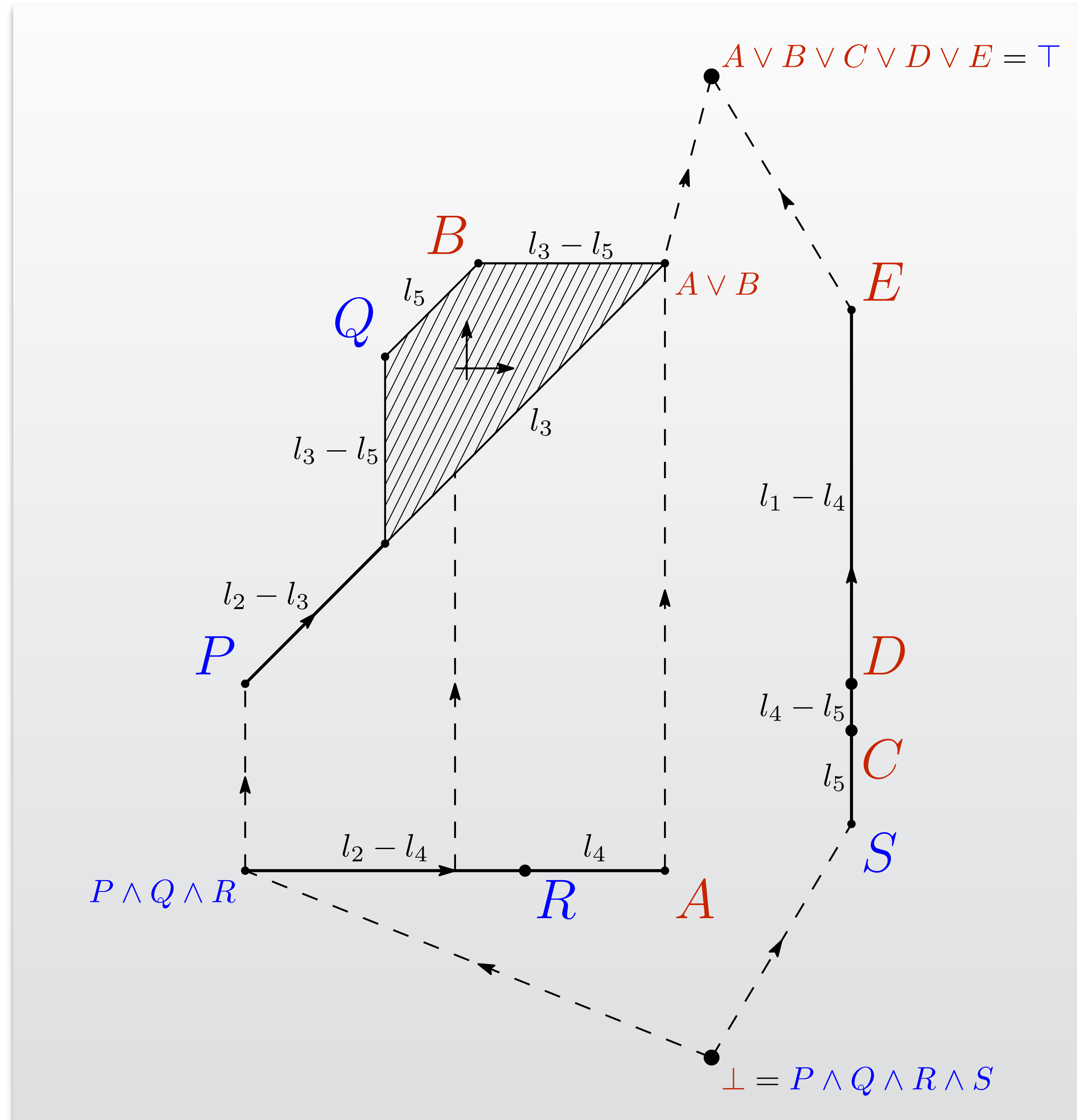
Quantitative Concept Analysis [Pavlovic, 2012]

- Lawvere metrized

	P	Q	R	S
Alice	★★	★★★	★★★★	
Bob	★★	★★★★★		
Carol				★★★★★
Dan				★★★★
Eve				★

$d(A, P) = l_2, d(A, Q) = l_3, \ \ldots$

where $0 \leq l_5 \leq \cdots \leq l_1 < l_0 = \infty$



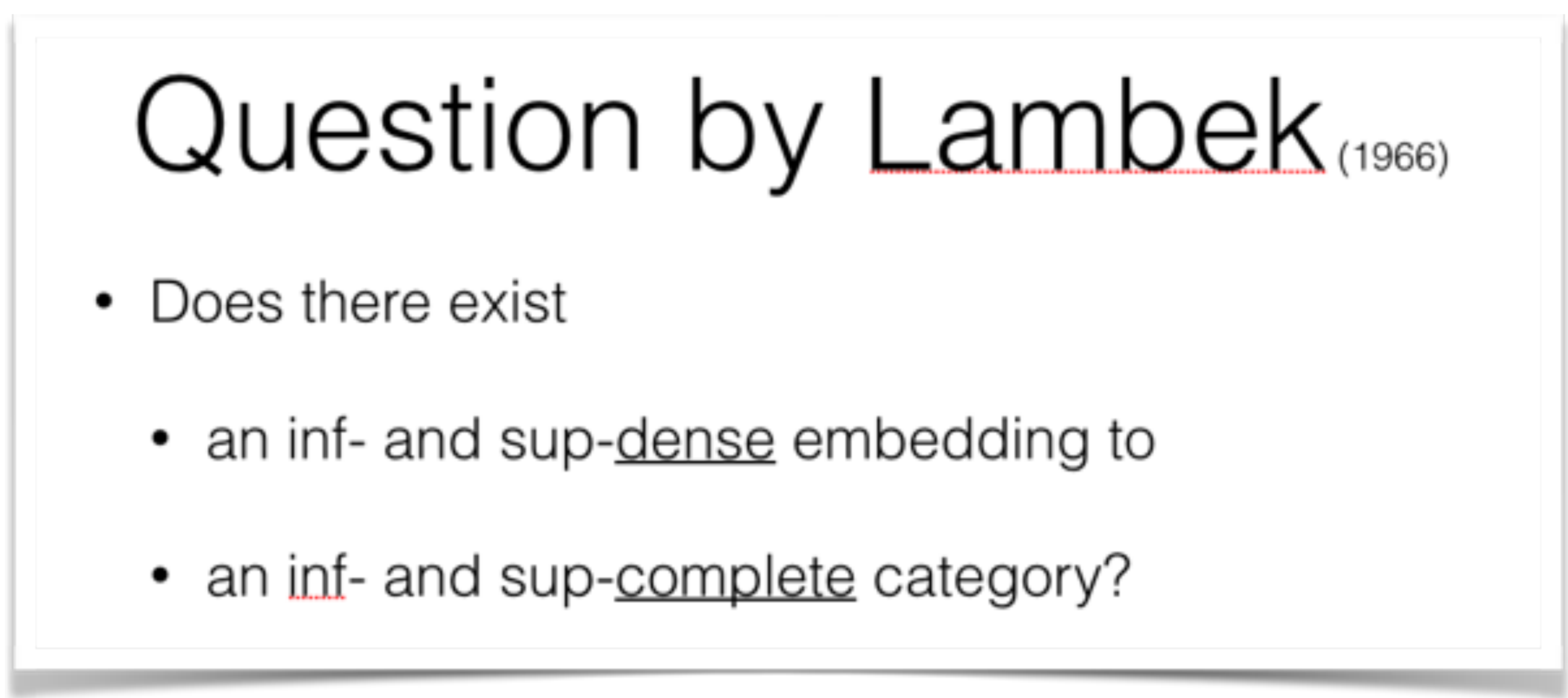
Bicompletions *of Categories*

Question by Lambek (1966)

- Does there exist
  - an inf- and sup-dense embedding to
  - an inf- and sup-complete category?

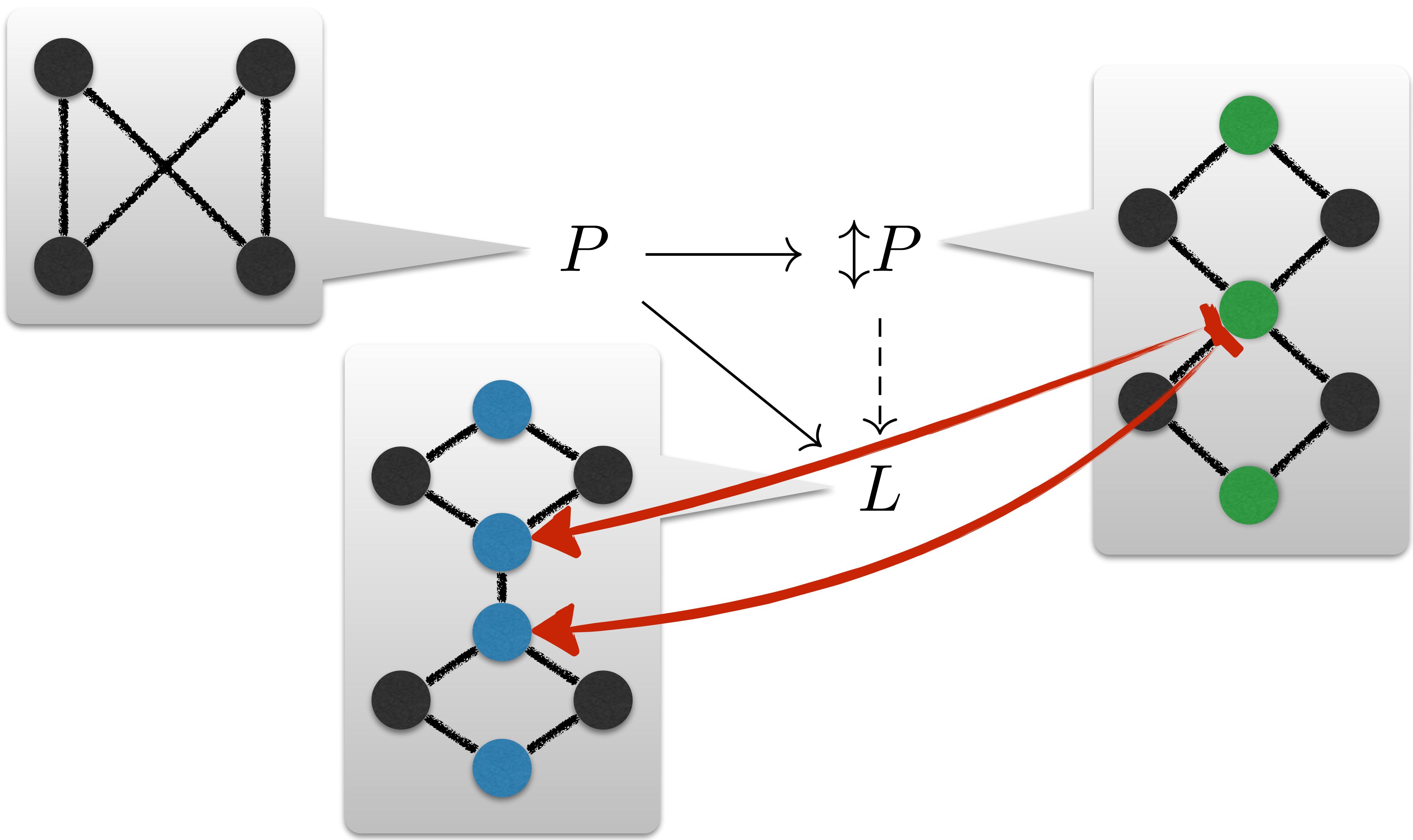
Answer by Isbell (1968)

- Any inf- and sup-dense embedding of the group $\mathbf{Z}_4$ is not inf-complete (nor sup-complete)

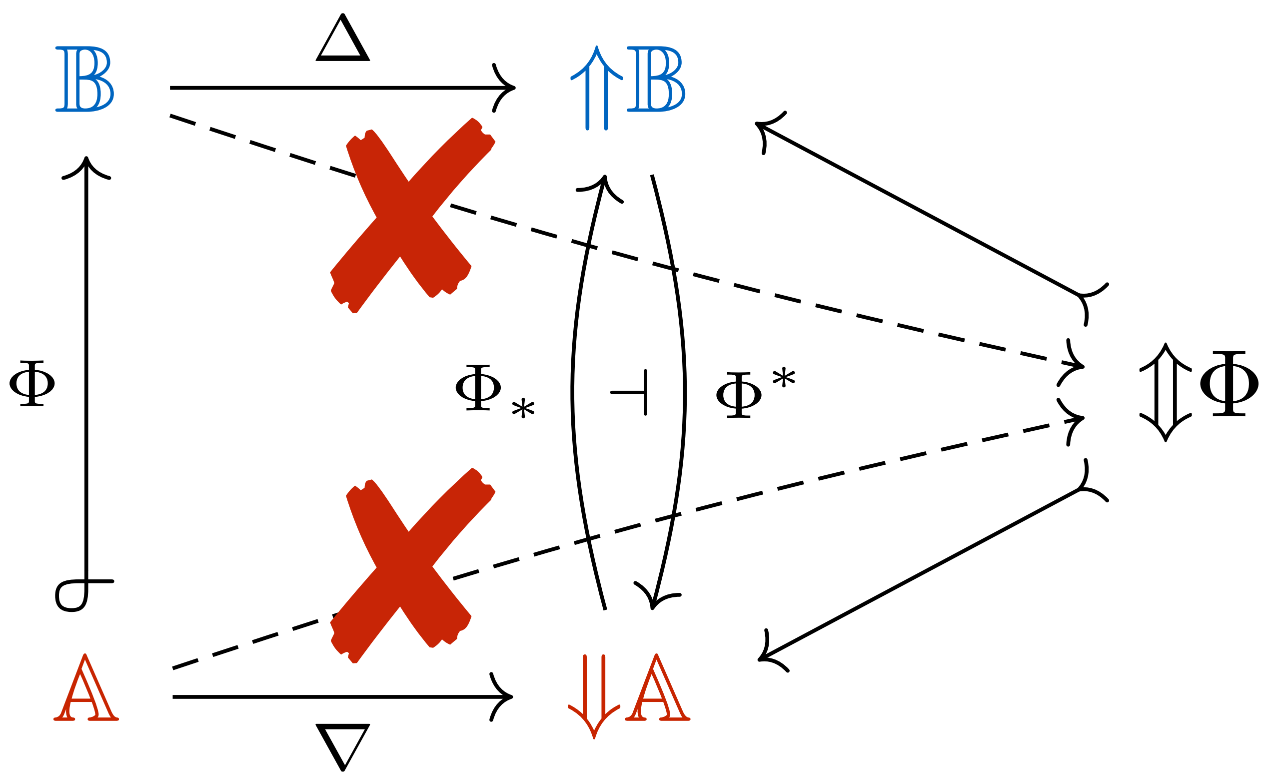


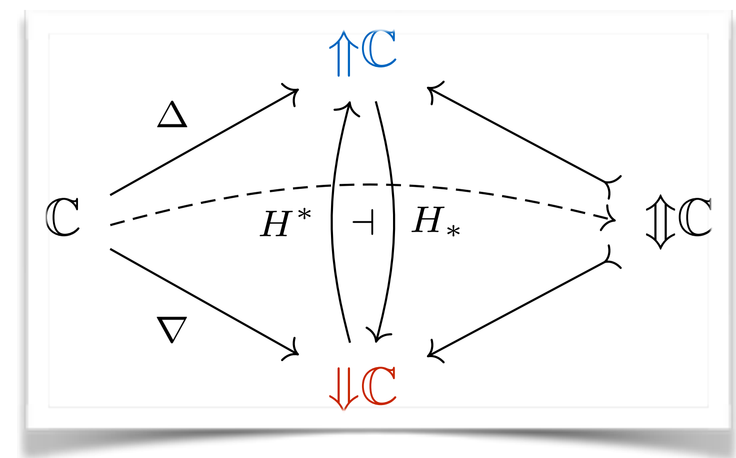
Two Universalities

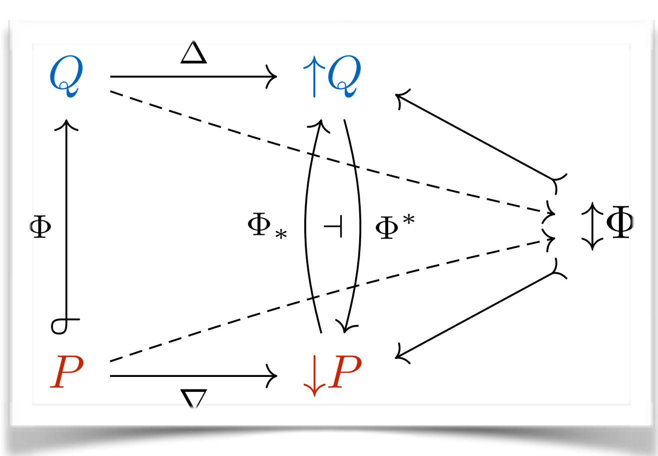
- Cannot be a single self-dual universality

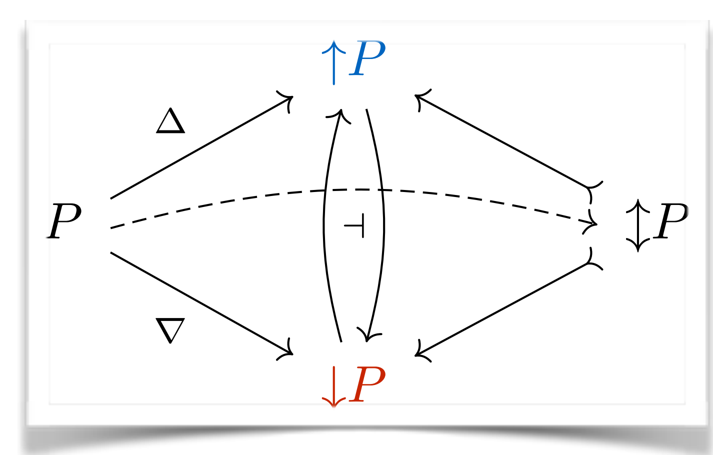


Trouble with $\mathcal{V}$ = Set

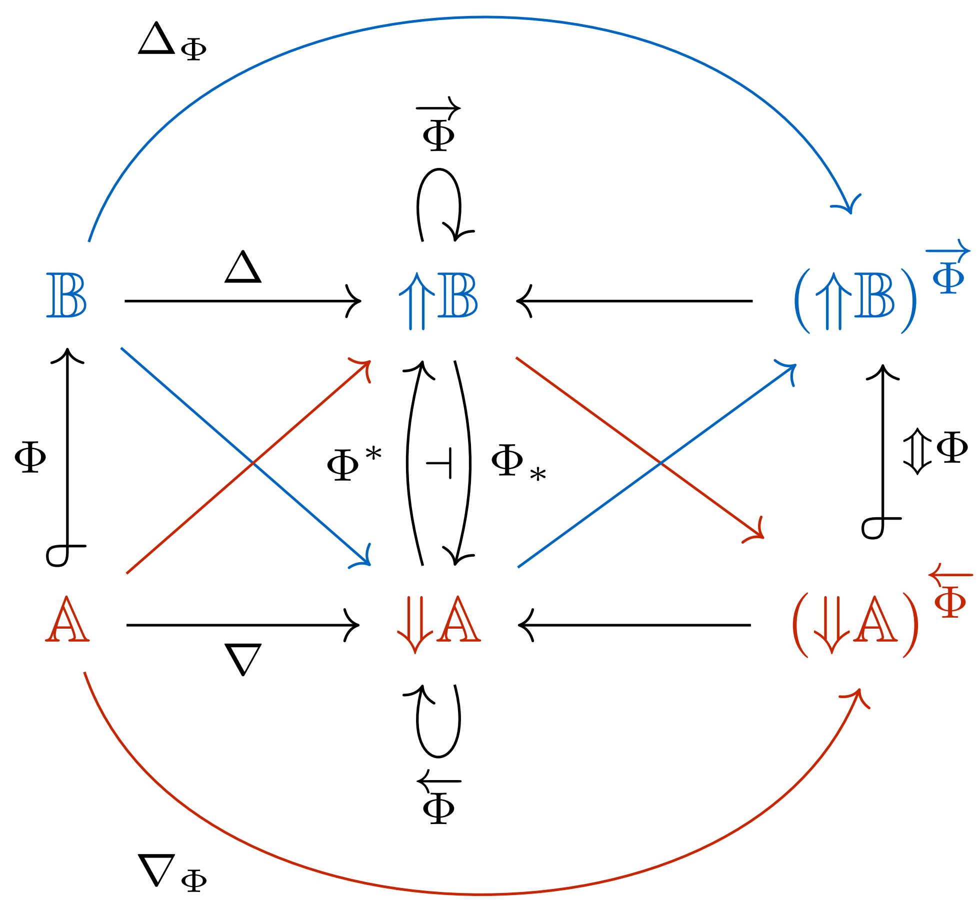


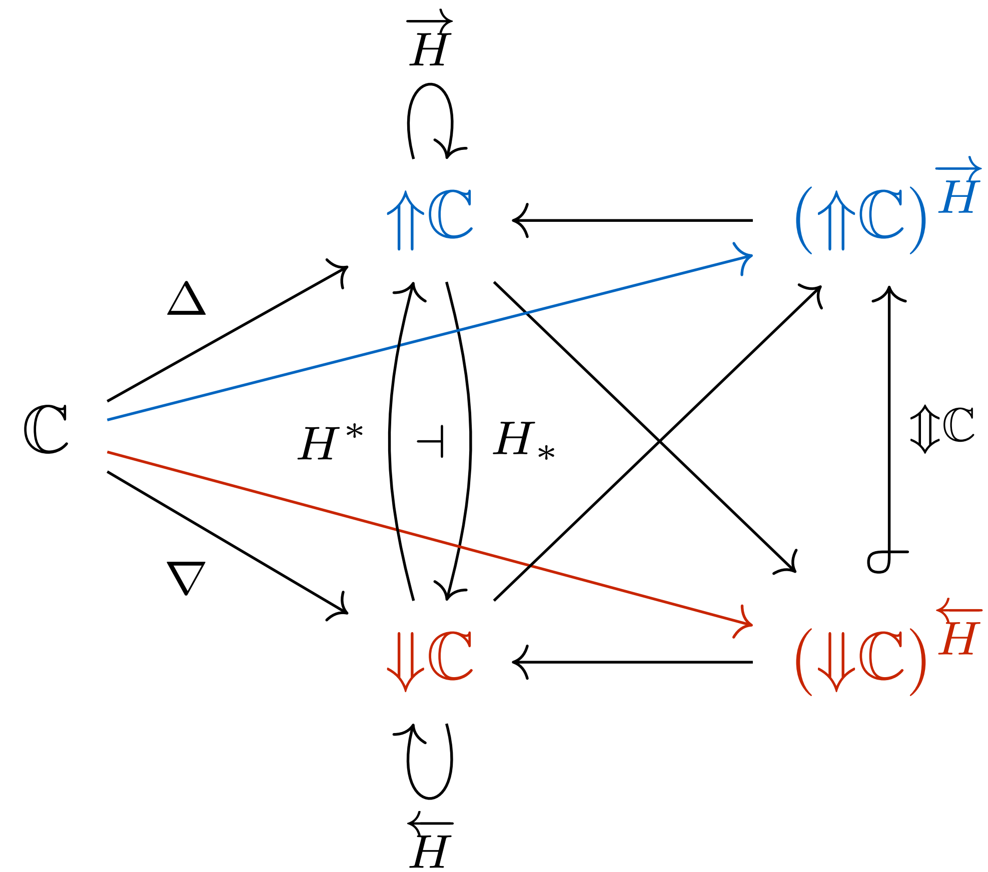


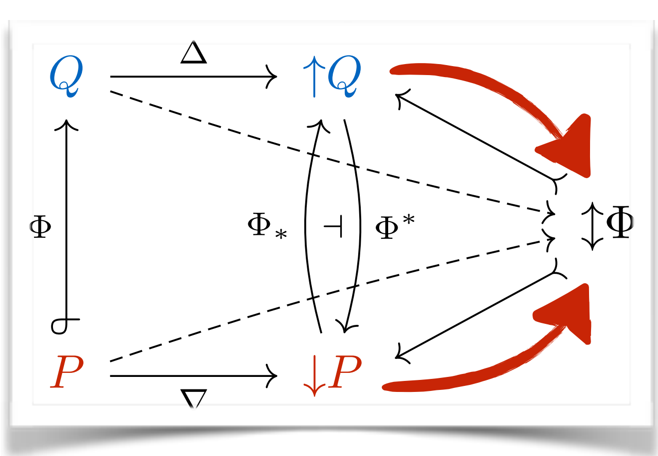


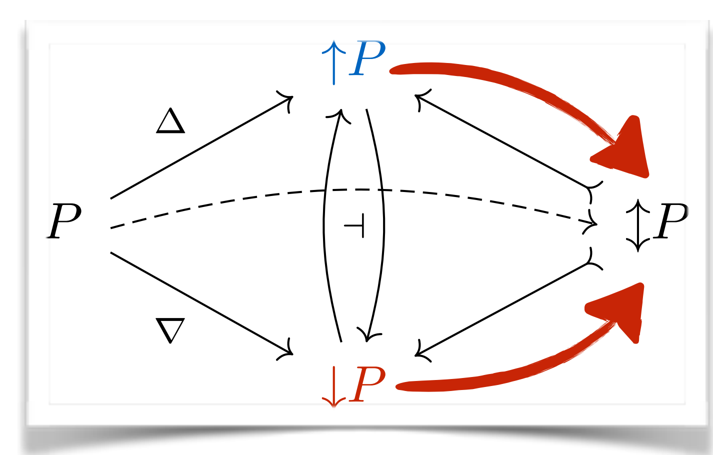


Our Proposal



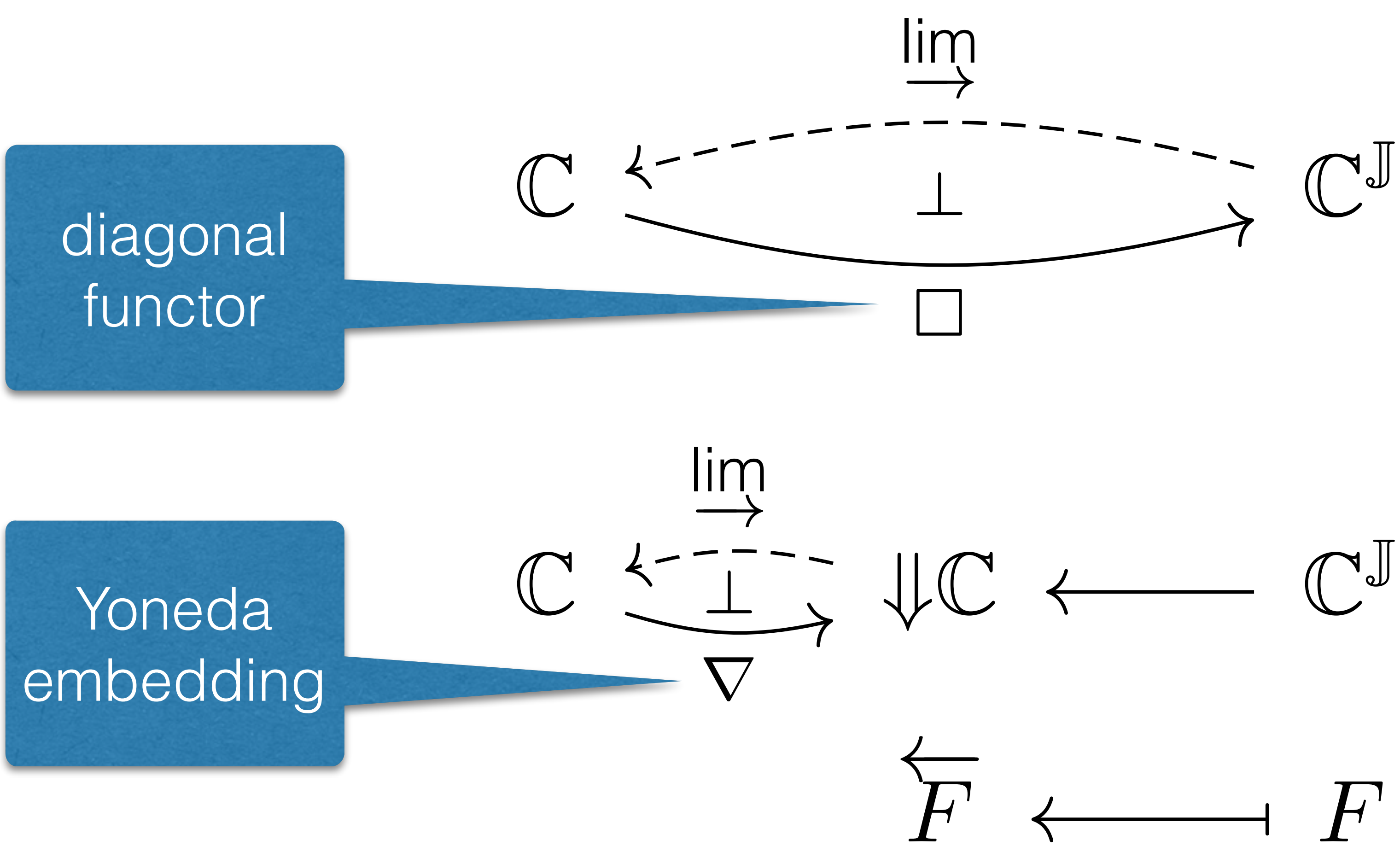






Supremum along Discrete Fibration

- $\varinjlim F = \varinjlim \overleftarrow{F}$ $(F : \mathbb{J} \to \mathbb{C},\ \overleftarrow{F} : \mathbb{C}^{\mathrm{op}} \to \mathrm{Set})$



Limit inferior

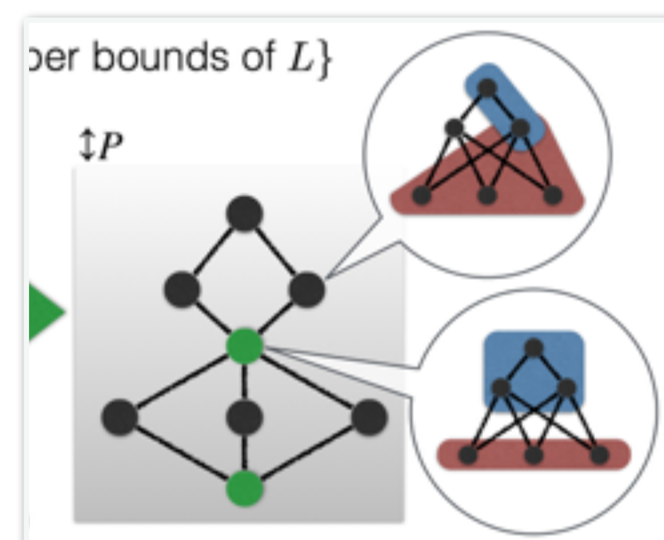
- The supremum of lower bounds

$$\overrightarrow{\lim} F := \underrightarrow{\lim}\, H^* \overrightarrow{F}$$

$$\mathbb{C} \xleftarrow{\quad \overrightarrow{\lim} \quad} \mathbb{C}^{\mathbb{J}}$$

$$\mathbb{C} \underset{\nabla}{\overset{\underrightarrow{\lim}}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} {\Downarrow}\mathbb{C} \xleftarrow{\ H^*\ } {\Uparrow}\mathbb{C} \longleftarrow \mathbb{C}^{\mathbb{J}}$$

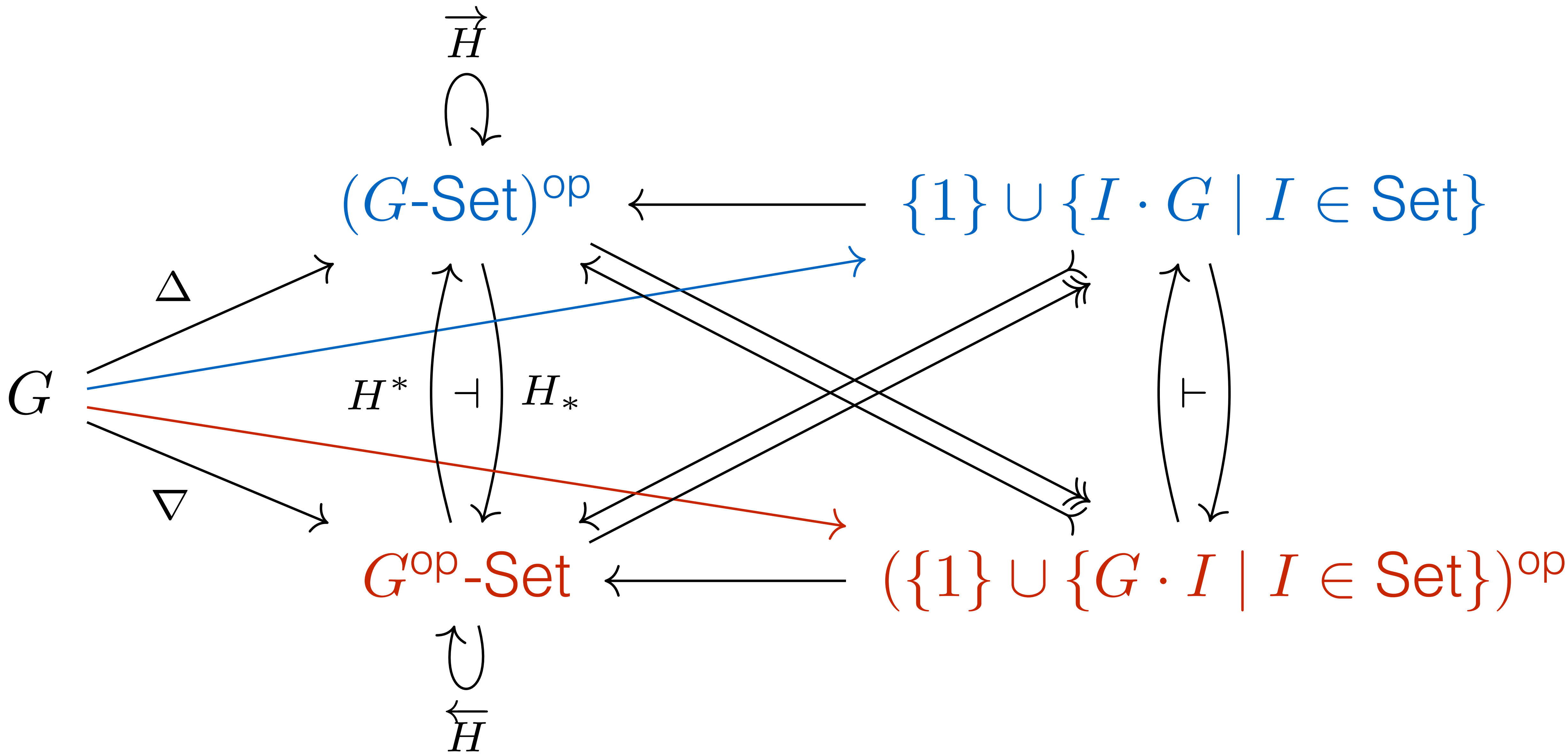
$$H^* \overrightarrow{F} \longleftarrow\!\shortmid\ \overrightarrow{F} \longleftarrow\!\shortmid\ F$$



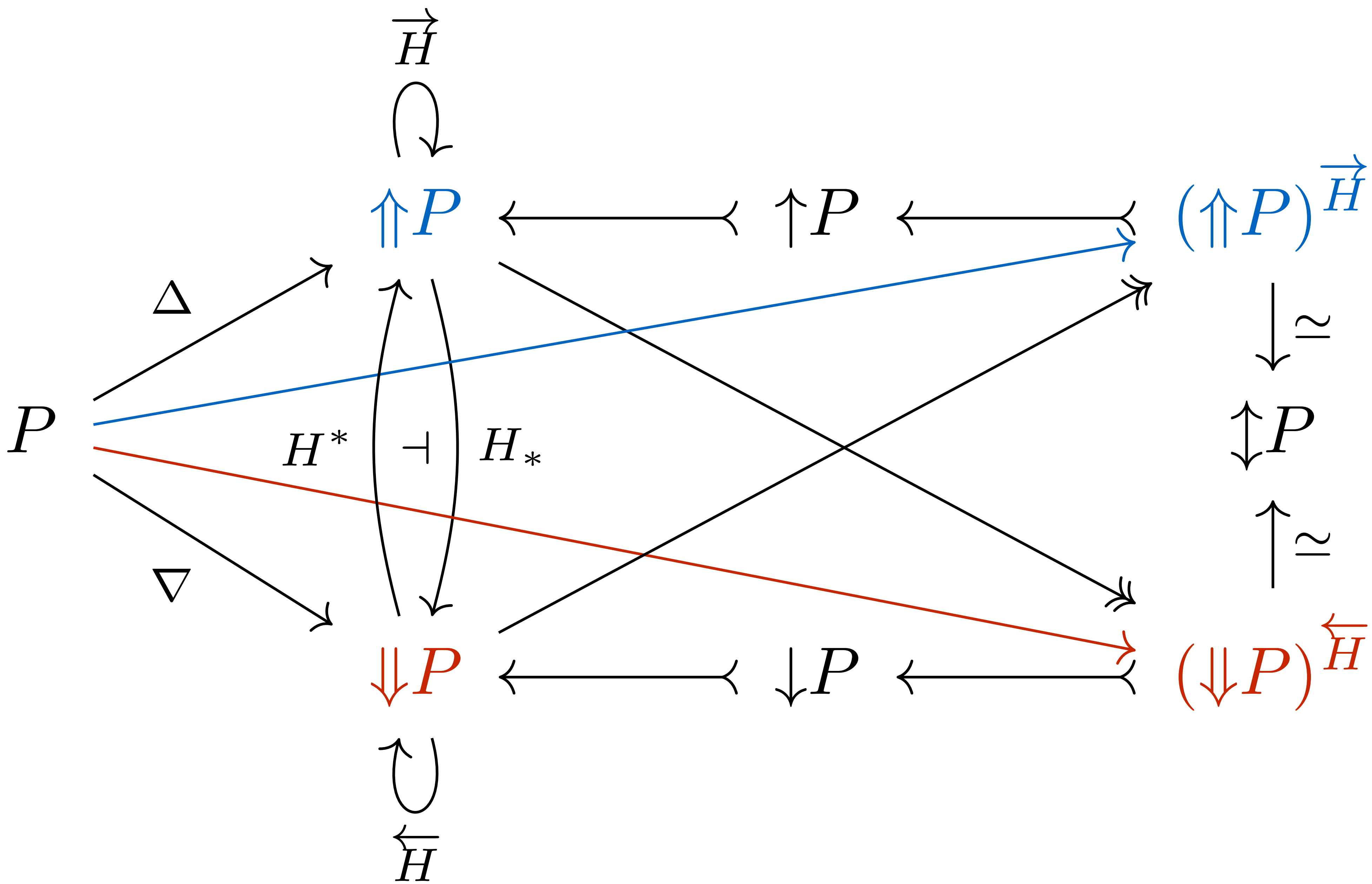
Theorems

- $(\Downarrow\mathbb{C})^{\overleftarrow{H}}$ is the free $\varinjlim$-completion of $\mathbb{C}$.
- $\mathbb{C} \to (\Downarrow\mathbb{C})^{\overleftarrow{H}}$ is $\varprojlim$- and $\varinjlim$-preserving.

Example: Completing Groups

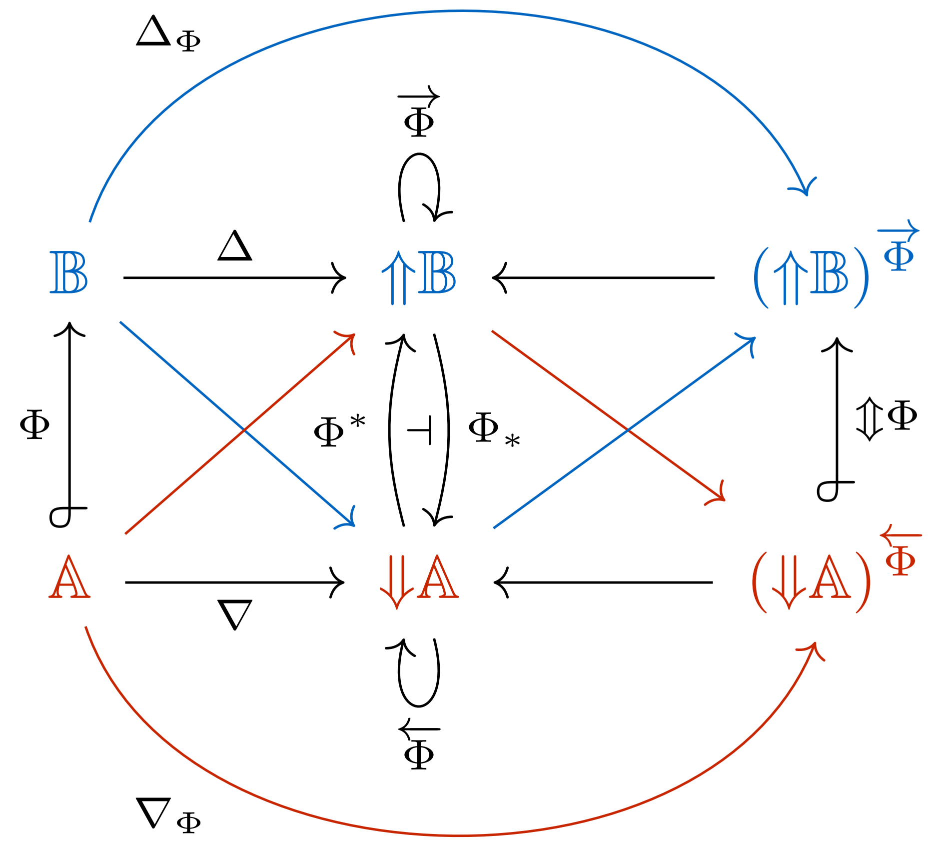


Example: Completing Posets



Summary

- Algebras, instead of fixed points
- Factorization (future work)
- Add limits inferior
    - Not always limits superior
- Minimality (future work)



References

- Isbell, John R. "Small subcategories and completeness." Theory of Computing Systems 2.1 (1968): 27-50.

- Lambek, Joachim. "Completions of Categories: Seminar lectures given 1966 in Zürich." Springer Berlin Heidelberg, 1966.

- Pavlovic, Dusko. "Quantitative concept analysis." Formal Concept Analysis. Springer Berlin Heidelberg, 2012. 260-277.

- Wille, Rudolf. "Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts." Ordered Sets. Springer Netherlands, 1982. 445-470.