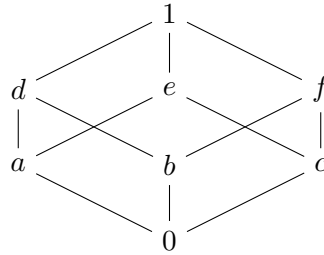


Everything I know so far about [Formal Concept Analysis]

Lucas Carr

May 20, 2024



Contents

| | | |
|----------|-------------------------------------|----------|
| 1 | Introduction | 2 |
| 1.1 | Lattices | 2 |
| 1.1.1 | Supremum and infimum | 2 |
| 1.2 | Formal Contexts | 2 |
| 1.3 | Formal Concepts | 3 |
| 1.4 | Concept Hierarchies | 3 |
| 1.5 | The Basic Theorem | 3 |
| 1.5.1 | I | 3 |
| 1.5.2 | II | 4 |
| 2 | Closure Systems | 6 |
| 2.1 | Defining a Closure System | 6 |
| 2.2 | Closure Operators | 6 |
| 3 | Implications | 7 |
| 3.1 | Precursor | 7 |

1 Introduction

1.1 Lattices

A lattice \mathcal{C} is a poset s.t. for any pair $(a, b) \in \mathcal{C}$, the supremum $a \wedge b$, and infimum $a \vee b$ exist. We extend this to a complete lattice, which has the requirement that for any subset $\mathcal{D} \subseteq \mathcal{C}$ the supremum $\bigvee \mathcal{D}$ and infimum $\bigwedge \mathcal{D}$ exist.

1.1.1 Supremum and infimum

The supremum of two elements is defined as the *least upper bound* of those two elements. We can obviously extend this over more than two elements. Given the set $S = \{1, 2, 3, 4, 5\}$, where $S \subset \mathbb{N}$. The supremum, $\bigvee S = 5$; similarly, the infimum $\bigwedge S = 1$.

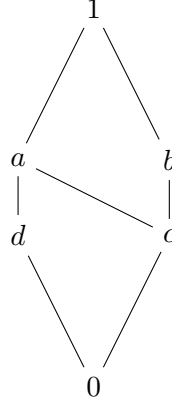


Figure 1: $d \vee c = a$, $b \wedge a = c$

When we are talking about lattices (or Hasse diagrams in general), we can refer to the supremum and infimum as the *meet* and *join*. Refer to 1.5 for more pertinent discussion regarding lattices.

We also have infimum (supremum) satisfying the following

- $x \wedge y = y \wedge x$ (commutativity)
- $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ (associativity)
- $x \wedge x = x$ (idempotency - but arguably, reflexivity)

We use these later on to show that any finite lattice is a complete lattice.

There are also the following interesting properties:

- $a \vee 0 = a$
- $a \leq b \implies a \vee b = b$ ¹

1.2 Formal Contexts

A Formal Context is a triple $\langle G, M, I \rangle$ where G refers to a set of objects, M to a set of properties, and I an incidence relation over $G \times M$.

We have derivation operators A' and B' ; for A' , where $A \subseteq G$, the derivation operator tells us which properties belong to the objects in A , the dual holds for properties and their objects. **Formally**,

Definition 1.1

$$A' := \{m \in M \mid \forall g \in A, gIm\}$$

$$B' := \{g \in G \mid \forall m \in B, gIm\}$$

¹This is useful because it enables to move from orders (posets) to lattices

We also have closure operators, A'' , which works intuitively by applying the derivation operator on A (B), which yields a set of properties. Then applying it again on A' (B'), which yields back a set of objects (properties).

Proposition 1.2 *For subsets $A, B \subseteq G$ (defined dually for properties $C, D \subseteq M$), we have*

- a. $A \subseteq B \implies B' \subseteq A'$
- b. $A \subseteq A''$
- c. $A' = A'''$

For more natural discussion, *a* describes the behaviour that if we have two sets of objects A and B , where $A \subseteq B$; then it follows that objects in A will have at *least* all the properties of objects in B .

1.3 Formal Concepts

Presume we are working with a formal context $\langle G, M, I \rangle$.

Definition 1.3 (A, B) is a **formal concept** of our formal context iff $A \subseteq G$, $B \subseteq M$, $A' = B$, and $B' = A$

A is called the **extent**, and B is called the **intent**. We can refer to the set of all formal concepts of a formal context as a $\mathcal{B}(G, M, I)$.

1.4 Concept Hierarchies

When we think about concepts, we typically think of them in a structure (something like a taxonomy, or ontology). That is, we think of *sub* and *super* concepts. For example, Dog is a subconcept of Mammal; and so is Cat. However, Dog and Cat are not sub nor superconcepts of one another (so, in some sense we have a partial order here). Formal Concepts have the same idea.

Definition 1.4 Let (A_1, B_1) and (A_2, B_2) be formal concepts of some $\mathcal{B}(G, M, I)$. We say that (A_1, B_1) is a **subconcept** of (A_2, B_2) (equivalently, (A_2, B_2) is a **superconcept** of (A_1, B_1)) and use the \leq sign to express this. Thus we have

$$(A_1, B_1) \leq (A_2, B_2) : \iff A_1 \subseteq A_2 \iff B_2 \subseteq B_1.$$

1.5 The Basic Theorem

1.5.1 I

The first part of the Basic Theorem tells us that given some formal concept (G, M, I) , the concept lattice of $\mathcal{B}(G, M, I)$ is a complete lattice. So, constructing a lattice from a formal context, will always result in a complete lattice.

From $\mathcal{B}(G, M, I)$, let $\{(A_t, B_t) | t \in T\} \subseteq \mathcal{B}(G, M, I)$ be some arbitrary subset of formal concepts. We can then define the supremum and infimum of this subset as:

$$\begin{aligned} \bigvee_{t \in T} (A_t, B_t) &:= \left(\left(\bigcup_{t \in T} A_t \right)', \left(\bigcap_{t \in T} B_t \right) \right) \\ \bigwedge_{t \in T} (A_t, B_t) &:= \left(\left(\bigcap_{t \in T} A_t \right), \left(\bigcup_{t \in T} B_t \right)' \right) \end{aligned}$$

Now, we want to show that there is some complete lattice \mathcal{V} that is isomorphic to the concept lattice given by $\mathcal{B}(G, M, I)$.

| | Fur | Feathers | Scales | Lives on Land | Lays Eggs | Cold-Blooded |
|------------|-----|----------|--------|---------------|-----------|--------------|
| Mammals | × | | | × | | |
| Birds | | × | | × | × | |
| Reptiles | | | × | × | × | × |
| Amphibians | | | × | × | × | × |
| Fish | | | × | | | × |

Suppose we are interested in Mammals and Birds, $\{Mammals\}' = \{Fur, LivesonLand\}$; $\{Mammals\}'' = \{Mammals\}$. $\{Birds\}' = \{Feathers, LivesonLand, LaysEggs\}$; $\{Birds\}'' = \{Birds\}$.

Using the basic theorem from above, we can determine the supremum and infimum of our set of concepts. We have the set of concepts from above, lets call this $S := (A_t, B_t)$. We can calculate $\bigvee S$ by

$$\begin{aligned} & \left(\bigcup_{t \in T} A_t \right)'', \left(\bigcap_{t \in T} B_t \right) \\ & \left(\{Mammals, Birds\}'', \{LivesonLand\} \right) \\ & \left(\{Mammals, Birds, Amphibians, Reptiles\}', \{LivesonLand\} \right) \end{aligned}$$

Dually, we can find the infimum $\bigwedge S$, given as: (which is the bottom element)

$$\begin{aligned} & \left(\bigcap_{t \in T} A_t, \left(\bigcup_{t \in T} B_t \right)'' \right) \\ & \left(\emptyset, (\{Fur, Feathers, LivesonLand, LaysEggs\})' \right) \\ & (\emptyset, \emptyset) \end{aligned}$$

The first part of the Basic Theorem, which we have just gone through, shows that for the concept lattice of a given formal context, $\mathcal{B}(G, M, I)$, we have an infimum and supremum for any arbitrary subset of concepts. This is just our definition of a complete lattice - so we have that any formal context has a complete lattice.

1.5.2 II

The second part of the basic theorem tells us that every complete lattice is *isomorphic* to a concept lattice. That is, given a complete lattice (\mathcal{L}, \leq) , we can find some $\mathcal{B}(G, M, I)$ that is isomorphic to \mathcal{L} .

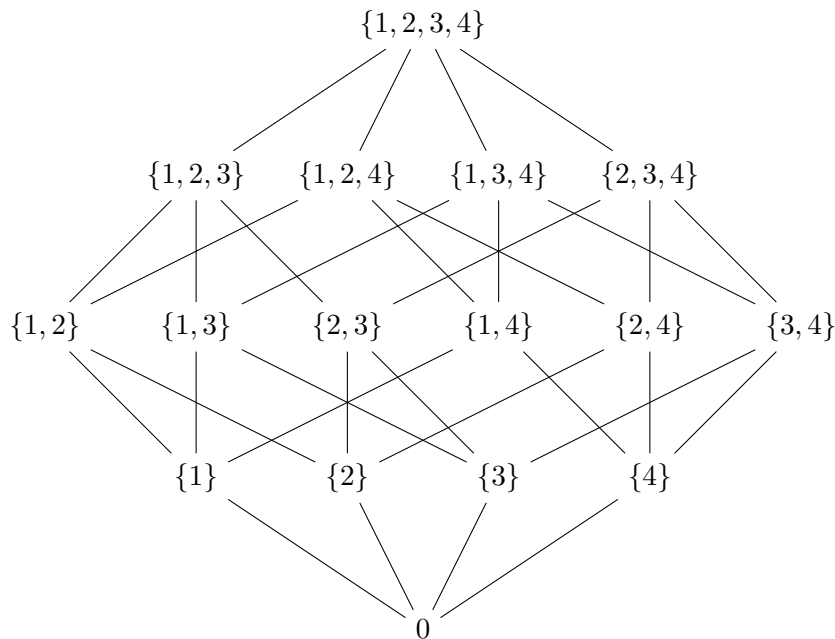


Figure 2: Example of a complete lattice of $(\{1, 2, 3, 4\}, \subseteq)$

Example 1.5

This is a complete lattice (it is finite, so it is complete); we can find a relatively simply formal context, which is isomorphic to the lattice: we need two sets, the set of supremum-irreducible elements, and the set of infimum-irreducible elements. For small lattices, we can do this via inspection - the supremum-irreducible (infimum-irreducible) elements are those which which only have a single line below (above) the element.

Thus, our supremum-irreducible set is $\{1, 2, 3, 4\}$;

and our infimum-irreducible set is $\{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}\}$.

The supremum-irreducible set represents the objects, and infimum-irreducible set represents the properties. The reason for this is intuitive: we start bottom-up with our maximally sub-concepts (remember, the distinction between properties and objects is a modelling problem), these basic elements are the ones which supremum-irreducible, i.e. they cannot be written as a super-concept of some other elements. Conversely, our infimum-irreducible elements are the most general (other than \top) elements of our lattice, they cannot be written as the sub-concepts of elements.²

Table 1: The formal context for 2

| | $\{1, 2, 3\}$ | $\{1, 2, 4\}$ | $\{1, 3, 4\}$ | $\{2, 3, 4\}$ |
|---|---------------|---------------|---------------|---------------|
| 1 | × | × | × | |
| 2 | × | × | | × |
| 3 | × | | × | × |
| 4 | | × | × | × |

[How do we determine the supremum(infimum)-irreducible sets without inspection?]

²We can of course write a infimum-irreducible element e as $e \vee \top$, but this is not useful.

2 Closure Systems

2.1 Defining a Closure System

If we are given a set M , a closure system \mathcal{C} on M is a set of subsets of M ($\mathcal{C} \subseteq \mathcal{P}(M)$) for which:

- $M \in \mathcal{C}$
- if $D \subseteq \mathcal{C}$ then, $\bigcap D \in \mathcal{C}$

We might observe that the first point is ensured by the second point; this is because $\bigcap \emptyset = M$ - if we don't have this property, we have a contradiction. We might also observe that a closure system has a natural ordering; specifically, by subsumption. That is, \mathcal{C} is a poset with (\mathcal{C}, \subseteq) .

2.2 Closure Operators

3 Implications

3.1 Precursor

An implication $A \rightarrow B$ holds in a formal context (G, M, I) if

1. $\forall g \in G, g' \models A \rightarrow B$
2. $A \rightarrow B \iff B \subseteq A''$
3. $A \rightarrow B \iff A' \subseteq B'$

(1) is obvious.

Looking at (2), we need to show that given $A \rightarrow B$, then $B \subseteq A''$. A' gives us the set of objects s.t. $\forall g \in A', A \subseteq g'$. A'' gives us the set of all properties these objects share. Since $A \rightarrow B$, and every object in A' has A , then every object in A'' should have B . Thus, $B \subseteq A''$.

In the other direction, suppose we have that $B \subseteq A''$, then we have either that $B \subseteq A$, in which case clearly $A \rightarrow B$. Or, we have that $\forall g \in A', g' \models A \rightarrow B$

For (3), if $A \rightarrow B$ then every object in A' must also have B in its attributes. And so, the objects described by A' are at most those described by B' .