

Rational Concept Analysis

Lucas Carr

A thesis submitted for the degree of
Master of Computer Science
Department of Computer Science
University of Cape Town

March 2025

Contents

Abstract	1
Acknowledgements	1
1 Mathematical Preliminaries	2
1.1 Order and Lattice Theory	2
1.1.1 Orders	2
1.1.2 Lattice Theory	4
1.2 Propositional Logic	5
1.2.1 Syntax	6
1.2.2 Semantics	6
1.2.3 Logical Consequence	7
1.2.4 Deduction Systems	8
2 Formal Concept Analysis	9
2.1 Basic Notions	9

Chapter 1

Mathematical Preliminaries

This chapter is intended to serve as a reference point, clarifying ideas and notation for the more fundamental concepts that will be used throughout the remainder of this dissertation. In the first section we discuss order and lattice theory. The third and final section introduces propositional logic, as well more general ideas in logic.

1.1 Order and Lattice Theory

This section refers extensively to the fundamental text by Davey and Priestley [1], as well as Ganter and Wille [2].

1.1.1 Orders

A *binary relation* R over the sets X and Y is a set of ordered pairs (x, y) where $x \in X$ and $y \in Y$. We may choose to express $(x, y) \in R$ using infix notation and write xRy , which tells us that R relates x to y .

Certain binary relations, satisfying specific properties, occur frequently enough to warrant their own denomination. One such relation, which will be used in almost every section of this dissertation, is called a *partial order*.

Definition 1. A *partial-order* is a binary relation $\leq \subseteq X \times X$ that satisfies the following properties:

$$\text{(Reflexivity)} \quad x \leq x \tag{1.1}$$

$$\text{(Antisymmetry)} \quad x \leq y \text{ and } y \leq x \text{ implies } x = y \tag{1.2}$$

$$\text{(Transitivity)} \quad x \leq y \text{ and } y \leq z \text{ implies } x \leq z \tag{1.3}$$

for all $x, y, z \in X$.

Frequently, ‘preference’ is used as a metonymy for an order, and so in this context “element x is preferred to y ” should be taken to mean that $(x, y) \in \leq$, or simply $x \leq y$.

We write $x \not\leq y$ to indicate that (x, y) is not in the relation, and $x < y$ for the case where $x \leq y$ and $x \neq y$. In the scenario where $x \not\leq y$ and $y \not\leq x$ —i.e., that x and y are incomparable—we write $x \parallel y$. From a partial-order we can quite easily induce the notion of a *strict partial-order*.

Definition 2. A *strict partial-order* is a binary relation $< \subseteq X \times X$ that satisfies:

$$\text{(Irreflexivity)} \quad x \not< x \quad (1.4)$$

$$\text{(Asymmetry)} \quad x < y \text{ implies } y \not< x \quad (1.5)$$

$$\text{(Transitivity)} \quad x < y \text{ and } y < z \text{ implies } x < z \quad (1.6)$$

for all $x, y, z \in X$.

An *ordered set* is a pair (X, \leq) with X being a set and \leq being an ordering on the elements of X . We make notation easier, and use \mathbf{X} to denote the pair; moreover, the order relation associated with \mathbf{X} may be written \leq_X in settings where ambiguity arises. If \mathbf{Y} is a subset of \mathbf{X} , then \mathbf{Y} inherits the order relation from \mathbf{X} ; and so, for $x, y \in \mathbf{Y}$, $x \leq_Y y$ if and only if $x \leq_X y$.

We can visualize ordered sets through the use of *Hasse diagrams*.

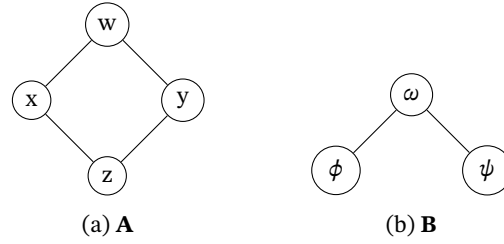


Figure 1.1: Hasse diagrams of two partially ordered sets

As an illustrative example, from the ordered set in Sub-figure 1.1a we read that $z \leq x$, as there is a (strictly) upward path from z to x . In fact, it is clear that $z \leq w, x, y, z$, or that “ z is preferred to every other element in \mathbf{A} ”. We say such an element is *minimal*.

More formally, an element $x \in \mathbf{X}$ is *minimal* with respect to the ordering if there exists no distinct element $y \in \mathbf{X}$ such that $y \leq x$. Conversely, we say x is *maximal* if there exists no distinct $y \in \mathbf{X}$ where $x \leq y$. Then x is the *minimum* element if $x \leq y$ for all $y \in \mathbf{X}$; the dual notion of a *maximum* is defined as we might expect.

Definition 3. Let \mathbf{X} and \mathbf{Y} be ordered sets with a mapping $\varphi : \mathbf{X} \rightarrow \mathbf{Y}$. We call φ an *order-preserving* (or, isotone) map if $x \leq_X y$ implies $\varphi(x) \leq_Y \varphi(y)$. It is an *order-embedding* if it is injective, and $x \leq_X y$ if and only if $\varphi(x) \leq_Y \varphi(y)$ for all $x, y \in X$. Finally, φ is an *order-isomorphism* if it is an order-embedding that is also *surjective*. The dual notion to an order-preserving map is an *order-reversing* (or, antitone) map.

1.1.2 Lattice Theory

Lattice theory studies partially ordered sets that behave well with respect to upper and lower bounds. Given an ordered set \mathbf{X} and a subset $\mathbf{Y} \subseteq \mathbf{X}$, the set of upper bounds of \mathbf{Y} is defined as

$$\mathbf{Y}^u := \{x \in \mathbf{X} \mid \forall y \in \mathbf{Y} : y \leq x\},$$

and the set of lower bounds, \mathbf{Y}^l , is defined dually. If \mathbf{Y}^u has a minimum element, then we call that element the *supremum* of \mathbf{Y} ; dually, if \mathbf{Y}^l has a maximum element, then we call that element the *infimum* of \mathbf{Y} . Instead of talking about the supremum of two elements $x, y \in \mathbf{X}$, we opt for the term *join* and write $x \vee y$, or $\bigvee \mathbf{Y}$ where $\mathbf{Y} \subseteq \mathbf{X}$. Instead of infimum, we say *meet* and write $x \wedge y$, or $\bigwedge \mathbf{Y}$.

A *lattice* is a non-empty partially ordered set \mathbf{L} in which every pair of elements $x, y \in \mathbf{L}$ for which $x \wedge y$ and $x \vee y$ exists. If $\bigwedge \mathbf{M}$ and $\bigvee \mathbf{M}$ exist for any subset $\mathbf{M} \subseteq \mathbf{L}$ then \mathbf{L} is a *complete lattice*.

1.1.2.1 Lattices as algebraic structures

From another perspective, a lattice can be seen as an algebraic structure where meet and join are binary operations on a set. Then, $\langle L, \vee, \wedge \rangle$ formally represents a lattice. Often, we simplify notation to just L when the context clearly indicates the operations.

If the set L is equipped with only one of these operations, the structure is called a *semi-lattice*. Specifically, the structure $\langle L, \vee \rangle$ (resp. $\langle L, \wedge \rangle$) is called a *join semilattice* (resp. a *meet semilattice*). Such a structure is a reduct of a lattice, satisfying the following properties for all $x, y, z \in L$:

$$\text{(associative)} \quad (x \vee y) \vee z = x \vee (y \vee z) \quad \text{(resp.)} \quad (x \wedge y) \wedge z = x \wedge (y \wedge z) \quad (1.7)$$

$$\text{(commutative)} \quad x \vee y = y \vee x \quad \text{(resp.)} \quad x \wedge y = y \wedge x \quad (1.8)$$

$$\text{(idempotent)} \quad x \vee x = x \quad \text{(resp.)} \quad x \wedge x = x \quad (1.9)$$

Finally, when considering both operations simultaneously, the interaction between join and meet operations within a lattice structure is governed by the absorption laws:

$$\text{(absorption)} \quad x \vee (x \wedge y) = x \quad \text{(resp.)} \quad x \wedge (x \vee y) = x \quad (1.10)$$

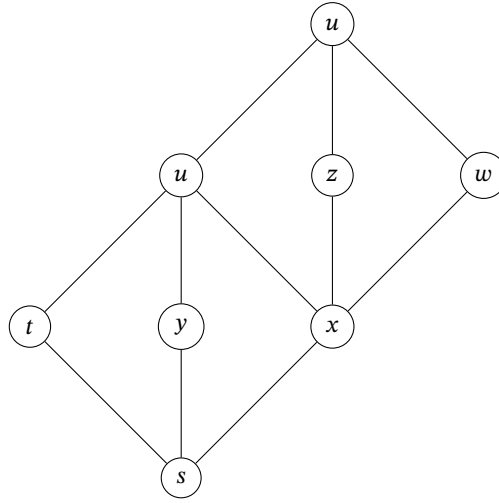


Figure 1.2: A complete lattice

Figure 1.2 illustrates a simple complete lattice; verifying that arbitrary meets and joins exist for all subsets. There is an obvious similarity between this representation of a lattice and the Hasse diagrams in Figure 1.1 used to describe partially ordered sets. This is perhaps obvious, given the opening remarks where we described a lattice as an ordered set, satisfying certain properties on upper and lower bounds. However, Lemma 1 makes the connection between the order and algebraic views of lattices concrete.

Lemma 1 (The Connecting Lemma). If \mathbf{L} is a lattice with $x, y \in \mathbf{L}$, then the following are equivalent:

1. $x \leq y$;
2. $x \vee y = y$;
3. $x \wedge y = x$.

Moving forward, we will no longer keep the order-theoretic and algebraic perspectives on lattices strictly separate; rather, we will tacitly adopt whichever viewpoint best suits the surrounding context. Lattices are central to Formal Concept Analysis and will be revisited in Chapter 2.

1.2 Propositional Logic

Propositional logic is a system for abstracting reasoning away from natural language. A *propositional statement* is a claim like “Tralfamadorians have one eye”; that is, something that can be assigned a value of *true* or *false* [3, p. 7]. More complex statements can be recursively constructed from simpler ones; our interest is then about which propositions follow from others.

1.2.1 Syntax

Propositional atoms are the construction blocks of a propositional language. They are indivisible statements that can either be true or false, and maybe be combined with the boolean connectives $\{\neg, \vee, \wedge, \rightarrow\}$ to construct more complex statements, or *formulae*. We denote propositional atoms with lower-case letters p, q, r, s , and t , and formulae with lower-case Greek letters $\alpha, \beta, \gamma, \phi, \psi$, and φ .

Not all combinations of atoms and boolean connectives result in meaningful expressions (for example, $\wedge \rightarrow \phi$) and so we make a distinction between any formulae and *well-formed formulae*. The latter being constructions that are valid with respect to the rules of the grammar defined in Backus-Naur form below. We are entirely uninterested in formulae that are not well-formed formulae, and so we drop the ‘well-formed’ suffix with the recognition that we shall never again mention the alternative [4, p. 33].

$$\phi ::= p \mid (\neg\phi) \mid (\phi_1 \vee \phi_2) \mid (\phi_1 \wedge \phi_2) \mid (\phi_1 \rightarrow \phi_2) \mid (\phi_1 \leftrightarrow \phi_2) \quad (1.11)$$

Any formula accepted by this grammar is said to be in the language \mathcal{L}^p , although we will drop the superscript where possible. To make reading this dissertation slightly more enjoyable, we may construct examples where we denote propositional atoms by monospaced text, enabling the expression of formulae like `tralfamadorians` \rightarrow \neg `human`, which should be interpreted as meaning “Tralfamadorians are not human”.

1.2.2 Semantics

In the previous section we described the syntax of propositional logic and how the boolean connectives enable the construction of arbitrary formulae from atoms. The semantics of propositional logic are concerned with how meaning can be ascribed to these formulae. The aim is to provide a method of answer questions like “when is this formula true?” or, “if ϕ, ψ, φ are true, what else must be true?”.

Propositional atoms were described as indivisible statements that can be assigned values of *true* or *false*. We now define a function that assigns a truth value to each proposition in the set \mathcal{P} of propositions, sometimes called the *universe of discourse*.

Definition 4. A *valuation* is a function $u : \mathcal{P} \rightarrow \{true, false\}$ that assigns a truth value to each propositional atom.

Given a set of atoms $\mathcal{P} = \{p, q, r\}$, we write \mathcal{U} to denote the set of all possible valuations. A valuation $u \in \mathcal{U}$ *satisfies* an atom $p \in \mathcal{P}$ if u maps p to *true*, then we write $u \models p$. Otherwise, we write $u \not\models p$ to indicate that u does not satisfy p (in this context, meaning u maps p to *false*) [3, p. 16]. We might represent the valuation that maps p and q to true but r to false by $\{p, q, \bar{r}\}$.

This satisfaction relation can be extended beyond propositional atoms to include more complex formulae, as described in Section 1.2.1. For any $\phi, \psi \in \mathcal{L}$

- $u \models \neg\phi$ if and only if $u \not\models \phi$
- $u \models \phi \vee \psi$ if and only if $u \models \phi$ or $u \models \psi$
- $u \models \phi \wedge \psi$ if and only if $u \models \phi$ and $u \models \psi$
- $u \models \phi \rightarrow \psi$ if and only if $u \models \neg\phi$ or $u \models \psi$
- $u \models \phi \leftrightarrow \psi$ if and only if $u \models \phi \rightarrow \psi$ and $u \models \psi \rightarrow \phi$

Definition 5. For a valuation $u \in \mathcal{U}$ and formula $\phi \in \mathcal{L}$ we call u a *model* of ϕ if and only if u satisfies ϕ . The set of models of ϕ is constructed by $\hat{\phi} := \{u \in \mathcal{U} \mid u \models \phi\}$.

Satisfiability can be extended to sets of formulae so that a valuation $u \in \mathcal{U}$ *satisfies* the set $\Phi := \{\phi_0, \dots, \phi_n\}$ when $u \models \phi_i$ for all $0 \leq i \leq n$, and we write $u \models \Phi$. If no such valuation exists then Φ is unsatisfiable [3, p. 31].

1.2.3 Logical Consequence

The introduction of models leads quite naturally into a discussion on the matter of *logical consequence*. This is an attempt to formalise the process of determining what should (logically) follow from a set of true propositional statements.

As an example, if we know that “*Billy lived in Slaughterhouse 5*” and that “*The inhabitants of Slaughterhouse 5 survived the bombing of Dresden*”, then we may find it sensible to infer that “*Billy survived the bombing of Dresden*”. Of course, this inference being *sensible*—under some common concept of consequence—gives little insight into how logical consequence may be formalised.

Tarski [5, p. 410] provides such insight through a semantic account of logical consequence,

Definition 6. Let Γ be a set of formulae and φ a formula in the language \mathcal{L} . We say that φ is a *logical consequence* of Γ , and write $\Gamma \models \varphi$, if and only if every model of Γ is a model of φ , or equivalently if $\hat{\Gamma} \subseteq \hat{\varphi}$.

Example 1. If we model the earlier scenario in propositional logic (where we initialise *Billy* with b , *slaughterhouse 5* with h , and *survived* with s), we can answer the question of whether $b \rightarrow s$ is a logical inference by determining if the models of $b \rightarrow s$ are a superset of the models of $\{b \rightarrow h, h \rightarrow s\}$. By the earlier satisfaction relation, we know that the implication $b \rightarrow s$ is satisfied by any valuation that satisfies either $\neg b$ or s . It is then quite simple to spot that the only possible valuations that are not models of this implication are $\{b, \bar{s}, h\}$ and $\{b, s, \bar{h}\}$ neither of which are models of $\{b \rightarrow h, h \rightarrow s\}$.

Before we progress any further, some attention needs to be given to the distinction between what we call the *object* and *meta-languages*. In Example 1, propositional logic forms the object-language; and so the boolean connectives are object-level connectives. Logical consequence, however, is a concept pertaining to the meta-language. If we consider the sentence $b \rightarrow s$, what this expression tells us is that in any valuation where b is true, s must also be true.

The meta-language facilitates reasoning over the elements in the object-language itself.

1.2.4 Deduction Systems

Chapter 2

Formal Concept Analysis

Formal Concept Analysis (FCA) provides a simple, and yet mathematically rigorous, framework for identifying and reasoning about “concepts” and their corresponding hierarchies in data [2, 6]. The central view of concepts as a dual between *extension*—what one refers to as instances of a concept—and *intension*—what meaning is ascribed to a concept—is supported by a rich philosophical backing.

2.1 Basic Notions

The universe of discourse in FCA is made-up of sets of *objects* and *attributes*. Objects are extensional, they are the things pointed to as instances of some more general concept. In turn, attributes construct the intensional component of a concept.

We collect these extensional and intensional building blocks in a structure called a *formal context*, which includes a binary relation that allows traversal from extension to intension, and vice versa.

Definition 7. A *formal context* $\mathbb{K} = (G, M, I)$ is a triple comprised of a set of objects G , a set of attributes M , and a binary relation $I \subseteq G \times M$ referred to as an ‘incidence’ relation. For an object-attribute pair $(g, m) \in I$ we might say that “object g has the attribute m ”.

A formal context in some sense describes an open-world interpretation, and so $(g, m) \notin I$ is not usually interpreted as saying that “object g has the negation of the attribute m ”.

Example 2. Finite formal contexts of a reasonable size can be described entirely by a tabular representation. Each object corresponds to a row, and each attribute to a column.

	closure	associativity	identity	divisibility	commutativity
magma	×				
semigroup	×	×			
monoid	×	×	×		
group	×	×	×	×	
abelian group	×	×	×	×	×
loop	×		×	×	
quasigroup	×			×	
groupoid		×	×	×	
category		×	×		
semicategory		×			

Figure 2.1: A formal context showing necessary properties of group-like structures.

The tabular representation of a formal context allows for easy identification of the set of attributes that a given object satisfies: one need only scan across the respective row in the table and note where an ‘x’ symbol appears. This set of attributes is called the *object intent*. The dual notion of an *attribute extent* can similarly be found by scanning down the column of a given attribute. The utility of this visual metaphor diminishes when, as is often the case, we consider larger sets of objects or attributes. Instead, we opt for a more formal approach to determining the intents and extents for (sets of) objects and attributes, respectively.

Definition 8. Given a formal context (G, M, I) , the *derivation operators* are two order-reversing maps $(\cdot)^\uparrow : 2^G \rightarrow 2^M$ and $(\cdot)^\downarrow : 2^M \rightarrow 2^G$ where the order is given by subset inclusion. Then, for any subsets $A \subseteq G$ and $B \subseteq M$,

$$A^\uparrow := \{m \in M \mid \forall g \in A, \langle g, m \rangle \in I\}$$

$$B^\downarrow := \{g \in G \mid \forall m \in B, \langle g, m \rangle \in I\}$$

The derivation operators provide a clear way of describing, for a given set $A \subseteq G$ of objects, the set of attributes which every object in A satisfies, denoted A^\uparrow . As an illustration, given the set of objects {semigroup, monoid} from Figure 2.1, its derivation would be {closure, associativity}. It is quite easy to spot that this is just the intersection of the object intents of semigroup and monoid.

Of course, these two functions can be composed; and so, $A^{\uparrow\downarrow}$ would yield the set of objects {semigroup, monoid, group, abelian group}, which can rather clumsily be described as “the set of all objects that satisfy all the attributes satisfied by semigroup and monoid”.

In fact, this double-application of derivation operators satisfies very specific properties,

$$(\text{monotonicity}) \quad A \subseteq A_1 \text{ implies } A^{\uparrow\downarrow} \subseteq A_1^{\uparrow\downarrow} \quad (2.1)$$

$$(\text{extensivity}) \quad A \subseteq A^{\uparrow\downarrow} \quad (2.2)$$

$$(\text{idempotency}) \quad A^{\uparrow\downarrow} = (A^{\uparrow\downarrow})^{\uparrow\downarrow} \quad (2.3)$$

for all $A, A_1 \subseteq G$. Thus, $(\cdot)^{\uparrow\downarrow}$ describes a closure operator on 2^G . The dual notion holds for attributes, and $(\cdot)^{\downarrow\uparrow}$ describes a closure operator on 2^M .

Proposition 1. Let (G, M, I) be a formal context with subsets $A_0, A_1, A_2 \subseteq G$ and $B_0, B_1, B_2 \subseteq M$ of attributes. Then,

- | | |
|--|--|
| 1. $A_0 \subseteq A_1 \Rightarrow A_1^{\uparrow} \subseteq A_0^{\uparrow}$ | 1. $B_0 \subseteq B_1 \Rightarrow B_1^{\downarrow} \subseteq B_0^{\downarrow}$ |
| 2. $A_0 \subseteq A_0^{\uparrow\downarrow}$ | 2. $B_0 \subseteq B_0^{\downarrow\uparrow}$ |
| 3. $A_0^{\uparrow} = A_0^{\uparrow\downarrow\uparrow}$ | 3. $B_0^{\downarrow} = B_0^{\downarrow\uparrow\downarrow}$ |

Definition 9. A *formal concept* of a formal context (G, M, I) is a pair (A, B) of subsets $A \subseteq G$ and $B \subseteq M$ that satisfies $A^{\uparrow} = B$ and $B^{\downarrow} = A$. Then, A is the concept *extent* and B is the *intent*. We write $\mathfrak{B}(G, M, I)$ to denote the set of all concepts of (G, M, I) .

Now explain why derivation of set of objects is always a concept intent by proposition 1. And the dual. and Galois connections

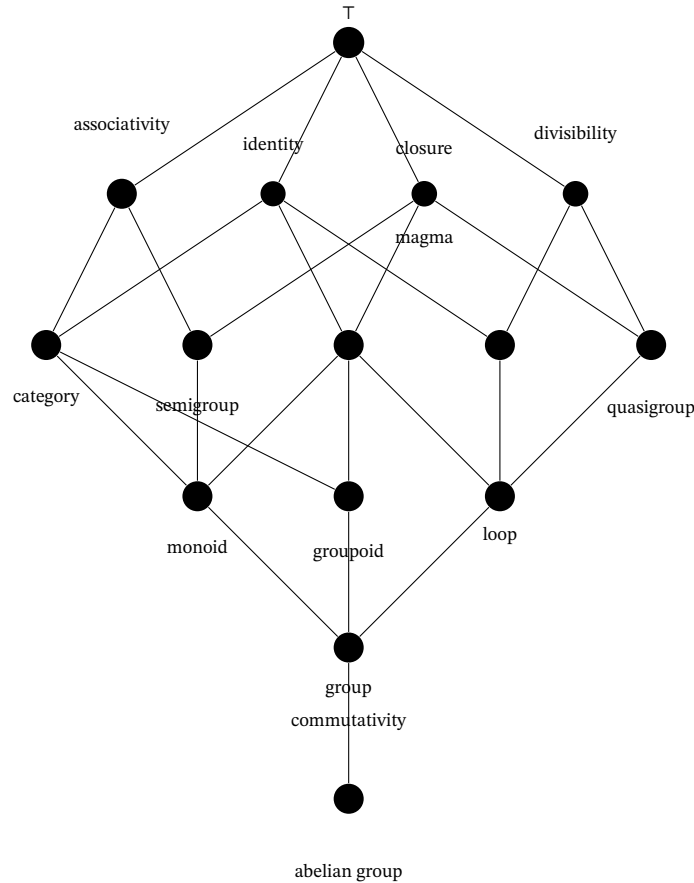


Figure 2.2: The concept lattice associated with the formal context in Figure 2.1

Bibliography

- [1] B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order*. Cambridge University Press, 2002.
- [2] B. Ganter and R. Wille, *Formal Concept Analysis: Mathematical Foundations*. Springer Berlin, Heidelberg, 1999.
- [3] M. Ben-Ari, *Mathematical logic for computer science*. Prentice Hall, 1993, vol. 13, pp. I–XI, 1–305.
- [4] M. Huth and M. Ryan, *Logic in Computer Science: Modelling and Reasoning about Systems*, 2nd ed. Cambridge University Press, 2004.
- [5] A. R. Turquette, A. Tarski, and J. H. Woodger, *Logic, Semantics, Metamathematics*. Oxford University Press (OUP), 1957, vol. 8, p. 87.
- [6] B. Ganter and S. Obiedkov, *Conceptual Exploration*. Springer Berlin Heidelberg, 2016, pp. 1–315.

Index

- binary relation, 2
- derivation operators
 - attribute extent, 10
 - object intent, 10
- extension, 9
- formal concept, 11
- formal concept analysis, 9
- formal context, 9
- Galois connections, 11
- Hasse diagrams, 3
- intension, 9
- join, 4
- lattice, 4
 - complete lattice, 4
- logical consequence, 7
- lower bound, 4
- maximal, 3
- maximum, 3
- meet, 4
- minimal, 3
- minimum, 3
- model, 7
- order-embedding, 3
- order-isomorphism, 3
- order-preserving map, 3
- order-reversing map, 3
- partial-order, 2
 - strict partial-order, 3
- propositional logic, 5
- satisfaction, 7
- upper bound, 4
- valuations, 6