The Galois Connection between Syntax and Semantics

Peter Smith
University of Cambridge

Contents

Preface

In his Dialectica paper 'Adjointness in foundations' (1969), F. William Lawvere writes of 'the familiar Galois connection between sets of axioms and classes of models, for a fixed [signature]'.

But even if long familiar folklore to category theorists, the idea doesn't in fact seem to be that widely known. The ideas here are pretty enough, elementary enough, and illuminating enough to be worth rehearsing briskly in an accessible stand-alone form.

Hence these brisk notes. If you are already happy with the idea of a 'poset' and ideas about mappings from one poset to another, proceed immediately to Chapter 2!

1
Partially ordered sets

This chapter introduces a class of structures – the 'posets' – which crop up in all sorts of contexts.

Of course, given that this type of structure is indeed exemplified all over the place, there won't be a great deal we can say about the general case. But still, our discussion will give us a chance to introduce some key ideas like the idea of an 'order-isomorphism'.

1.1 Posets introduced

Definition 1.1.1 *A* partially ordered set $\mathscr{P} = \langle P, \preccurlyeq \rangle$ *– henceforth, a* poset *– is a set P, carrying an ordering relation $\preccurlyeq$ which is reflexive, anti-symmetric and transitive. That is to say, for all $x, y, z \in P$,*

(i) *$x \preccurlyeq x$;*
(ii) *if $x \preccurlyeq y$ and $y \preccurlyeq x$, then $x = y$;*
(iii) *if $x \preccurlyeq y$ and $y \preccurlyeq z$, then $x \preccurlyeq z$.*

Why talk here of a 'partial' order? Because there is no requirement that a given element of the poset $\mathscr{P}$ be order-related to *any* element other than itself. But do note equally that our definition doesn't actually preclude a poset's ordering relation $\preccurlyeq$ being *total*: that is to say, it *may* be the case that for every x, y in the carrier set, either $x \preccurlyeq y$ or $y \preccurlyeq x$. Perhaps we should really call our orders 'at least partial': but plain 'partial' is absolutely standard.

Let's immediately draw out a simple consequence of this definition, one which will help to fix ideas about the kind of ordering we are talking about here:

Theorem 1.1.2 *A poset contains no cycles.*

By a 'cycle', we mean a chain of elements $p_0, p_1, p_2, \ldots, p_\alpha \in P$, *all different*, such that $p_0 \preccurlyeq p_1 \preccurlyeq p_2 \preccurlyeq \ldots \preccurlyeq p_\alpha \preccurlyeq p_0$ (to use an obvious notational shorthand).

Proof Suppose there is such a cycle. By transitivity, $p_0 \preccurlyeq p_1$ and $p_1 \preccurlyeq p_2$ implies $p_0 \preccurlyeq p_2$. By transitivity again, $p_0 \preccurlyeq p_2$ together with $p_2 \preccurlyeq p_3$ implies $p_0 \preccurlyeq p_3$. Keep on going in the same way until you've derived $p_0 \preccurlyeq p_\alpha$ (if α is infinite, our informal argument here is infinitary!). But we are already given $p_\alpha \preccurlyeq p_0$. So by anti-symmetry we can conclude $p_\alpha = p_0$, contradicting the assumption that the elements of the cycle are all distinct. $\square$

To put this graphically, suppose we draw a diagram of a poset. Put a dot for each distinct element, and a directed arrow from the dot corresponding to p to the dot corresponding to p' just when $p \preccurlyeq p'$. Then, because of transitivity, dots connected by a chain of arrows will also be directly connected. But the diagram will have no loops where we can follow a linked chain of directed arrows around and end up where we started – except for the trivial self-loops that correspond to the truth of each $p \preccurlyeq p$.

Here, then, are just a few examples of posets, starting with some based on familiar number systems before broadening out to give some non-numerical examples.

(1) The natural numbers $\mathbb{N}$ with $\preccurlyeq$ read as the natural 'less than or equals to' order relation $\leqslant$.
(2) $\mathbb{N}$ with the converse ordering – i.e. we take the order relation to be 'greater than or equals to'.
(3) $\mathbb{N}$ with the reordering $\preccurlyeq$ defined as follows in terms of the natural ordering $\leqslant$:
   (a) if m is even and n is odd, then $m \preccurlyeq n$;
   (b) if m and n are both odd, then $m \preccurlyeq n$ iff $m \leqslant n$;
   (c) if m and both n are both even, then $m \preccurlyeq n$ iff $m \leqslant n$.

   In other words, we put all the even numbers in their natural order before all the odds in *their* natural order.
(4) $\langle \mathbb{N}, = \rangle$ is also a poset, since identity is trivially the limiting case of a partial ordering (the 'smallest' partial ordering on any set, where each element is order-related only to itself).

(5) $\mathbb{N} \setminus \{0\}$ ordered by divisibility: so we put $m \preccurlyeq n$ iff m divides into n without remainder. (Here $\mathbb{N} \setminus \{0\}$ is of course the set of natural numbers with 0 removed.)
(6) $\mathbb{N} \setminus \{0, 1\}$ ordered by the same divisibility relation.
(7) The real numbers $\mathbb{R}$ with the natural 'less than or equals to' ordering.
(8) $\mathbb{R}$ with the converse ordering.
(9) The set $\mathbb{R}^2$ of ordered pairs of reals $\langle r, s\rangle$, taken with the 'lexicographic' ordering, i.e. putting $\langle r, s\rangle \preccurlyeq \langle r', s'\rangle$ just if either $r < r'$ or $r = r' \wedge s \leqslant s'$.
(10) Any set Σ of functions $f\colon \mathbb{R} \to \mathbb{R}$, putting $f \preccurlyeq g$ iff for all $x \in \mathbb{R}$, $f(x) \leqslant g(x)$.
(11) Choose an interpreted formal language L. Let $|\alpha|$ be the equivalence class containing all the L-wffs logically equivalent to α, and let E be the set of all such equivalence classes. Let $\Rightarrow$ be the relation that holds between the equivalence classes $|\alpha|, |\beta| \in E$ if $\alpha \vDash \beta$. Then it is easy to check that $\langle E, \Rightarrow\rangle$ is a poset. (If L comes equipped with a proof-system, then we could similarly define a poset in terms of a syntactic consequence relation rather than semantic consequence.)
(12) A downward-branching binary tree has a top node, and every node has zero, one or two children (think of 'truth-trees'). Take N to be set of nodes. Let's say that node n' is a *descendant* of n if n' is a child of n, or is a child of a child of n, or is a child of a child of a child .... Put $n \preccurlyeq n'$ iff $n' = n$ or n' is a descendant of n (in other words, n is on the same branch of the tree as n' and n is at least as high on the branch as n'). Then $\langle N, \preccurlyeq\rangle$ is a poset.
(13) Just for modal logicians: let W be a set of possible worlds in a Kripke frame for intuitionistic logic, and let $\sqsubseteq$ be the accessibility relation between worlds. Then $\langle W, \sqsubseteq\rangle$ is a poset.

And so it goes. Posets are ubiquitous.

We'll mention just two more cases for the moment:

(14) The powerset $\mathcal{P}(A)$ of some set A, with the members of $\mathcal{P}(A)$ ordered by set-inclusion $\subseteq$, is a poset.
(15) In fact, take *any* set $M \subseteq \mathcal{P}(A)$, i.e. take any arbitrary collection of subsets of a set A. Then $\langle M, \subseteq\rangle$ too is a poset.

Call these *inclusion posets.* A bit later, Theorem 1.6.1 will tell us that inclusion posets are in a good sense typical of all posets.

Note immediately that some of our example posets in fact have total orderings, while others have incomparable elements, i.e. elements x, y in the carrier set such that neither $x \preccurlyeq y$ nor $y \preccurlyeq x$. Exercise: which examples are which?

Note too the elementary point illustrated by the relationship between examples 5 and 6, and by the relationship between examples 14 and 15. Having worked out that the divisibility relation indeed gives us a partial ordering $\preccurlyeq$ on the naturals starting from 1, we don't have to do *more* work to show that the restriction of that relation to the numbers from 2 onwards is also a partial ordering. Likewise, having worked out that $\subseteq$ is a partial order on the carrier set $\mathcal{P}(A)$, we immediately see it remains a partial order when restricted to some $M \subseteq \mathcal{P}(A)$. This observation evidently generalizes to give us a trivial mini-theorem:

Theorem 1.1.3 *If $\mathscr{P} = \langle P, \preccurlyeq \rangle$ is a poset then so is $\mathscr{P}' = \langle P', \preccurlyeq' \rangle$ for any $P' \subseteq P$, where $\preccurlyeq'$ is the relation $\preccurlyeq$ restricted to P'.*

In such a case, we can say that $\mathscr{P}'$ is a *sub-poset* of $\mathscr{P}$.

A further, but equally elementary, point is illustrated by the relationship between examples 1 and 2, and between 7 and 8. This too we could dignify as a mini-theorem (whose proof is another trivial exercise):

Theorem 1.1.4 *If $\mathscr{P} = \langle P, \preccurlyeq \rangle$ is a poset then so is $\mathscr{P}^{op} = \langle P, \preccurlyeq^{op} \rangle$, where $\preccurlyeq^{op}$ is the opposite or converse of $\preccurlyeq$, i.e. for all $x, y \in P$, $x \preccurlyeq^{op} y$ iff $y \preccurlyeq x$.*

We'll say that $\mathscr{P}^{op}$ here is the *order-dual* of $\mathscr{P}$.

Of course, if a poset's order relation is naturally represented by e.g. '$\leqslant$', we'll want to represent the converse relation by '$\geqslant$': so the order dual of $\langle \mathbb{R}, \leqslant \rangle$ is $\langle \mathbb{R}, \geqslant \rangle$. Likewise, if a poset's order relation is naturally represented by '$\subseteq$', the converse relation which orders its dual will be represented by '$\supseteq$'. We'll use this sort of obvious notation in future.

1.2 Partial orders and strict orders

What we've just defined are sometimes also called *weakly ordered* sets. By contrast,

Definition 1.2.1 *A* strictly ordered *set $\mathscr{Q} = \langle Q, \prec \rangle$ is a set Q, carrying a relation $\prec$ which is irreflexive, asymmetric and transitive. That is to say, for all $x, y, z \in Q$,*

(i) $x \nprec x$;
(ii) *if* $x \prec y$ *then* $y \nprec x$;
(iii) *if* $x \prec y$ *and* $y \prec z$, *then* $x \prec z$.

For example, both $\langle \mathbb{N}, < \rangle$ and $\langle \mathbb{R}, < \rangle$, with '$<$' interpreted in the natural way both times, are *strictly* ordered sets. So is $\langle \mathbb{N} \setminus \{0\}, \prec \rangle$ where $m \prec n$ iff $m \neq n$ and m divides n without remainder. That latter example emphasizes that a *strict* order need not be a *total* order (so you need to watch the jargon here).

Should we care about developing a theory of strictly ordered sets as well as a theory of partially ordered sets? Yes and no. For note we have the following three-part theorem, whose proof can again be left as an exercise.

Theorem 1.2.2 *(i) Suppose* $\mathscr{P} = \langle P, \preccurlyeq \rangle$ *is a partially ordered set, and for* $x, y \in P$ *put* $x \prec y =_{\text{def}} x \preccurlyeq y \wedge x \neq y$. *Then* $\mathscr{P}^- = \langle P, \prec \rangle$ *is a strictly ordered set.*

(ii) Conversely, suppose $\mathscr{Q} = \langle Q, \prec \rangle$ *is a strictly ordered set, and for* $x, y \in Q$ *put* $x \preccurlyeq y =_{\text{def}} x \prec y \vee x = y$. *Then* $\mathscr{Q}^+ = \langle Q, \preccurlyeq \rangle$ *is a partially ordered set.*

(iii) The result of going from a partially ordered set to its strictly ordered correlate then back again returns us to the original structure: i.e. $(\mathscr{P}^-)^+ = \mathscr{P}$. *Similarly* $(\mathscr{Q}^+)^- = \mathscr{Q}$.

When we have a partially ordered set we can therefore immediately define a counterpart strictly ordered set (by ignoring identities); and likewise when we have a strictly ordered set we can define a counterpart partially ordered set (by adding identities). So, yes, if we are interested in partially ordered sets at all, we'll probably care equally about their strictly ordered counterparts. But no, because of the simple definitional link, we don't really need a separate theory to treat the strictly ordered sets. We can develop the theory of partially and strictly ordered sets treating either notion as the more basic and the other as derived. It is conventional and convenient to take the idea of posets as the fundamental one.

Henceforth, it will occasionally be notationally convenient to use $x \prec y$ as an abbreviation for $x \preccurlyeq y \wedge x \neq y$ (and likewise, $x \sqsubset y$ as an abbreviation for $x \sqsubseteq y \wedge x \neq y$, etc.).

1.3 Maps between posets

Next, we'll define three different, increasingly constrained, types of map that can relate two posets. Here's the first pair of notions we want:

Definitions 1.3.1 *Suppose that $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ are two posets. And let $f \colon P \to Q$ be a map between their carrier sets. Then*

(i) *f is* monotone *iff, for all $p, p' \in P$, if $p \preccurlyeq p'$ then $f(p) \sqsubseteq f(p')$;*
(ii) *f is an* order-embedding *iff for all $p, p' \in P$, $p \preccurlyeq p'$ iff $f(p) \sqsubseteq f(p')$.*

So a monotone mapping f preserves partial orderings. But note, it can – so to speak – 'squeeze them up'. In other words, if $p \prec p'$ we won't get the reversal $f(p) \sqsupset f(p')$, but we might have $f(p) = f(p')$.

An order-embedding however preserves more: it puts a faithful copy of $\mathscr{P}$'s ordering inside $\mathscr{Q}$.

Let's make that last point more precise, via a theorem, a definition, and another theorem. First, then, we remark that:

Theorem 1.3.2 *An order-embedding $f \colon P \to Q$ between $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ is an injective (one-one) map.*

Proof Suppose $f(p) = f(p')$. Then, by the reflexivity of $\sqsubseteq$, we have both $f(p) \sqsubseteq f(p')$ and $f(p') \sqsubseteq f(p)$. Since f is an order-embedding, it follows by definition that $p \preccurlyeq p'$ and $p' \preccurlyeq p$. And since $\preccurlyeq$ is anti-symmetric, it follows that $p = p'$.

So, contraposing, if $p \neq p'$ then $f(p) \neq f(p')$, so f is one-one or injective. □

That theorem prompts a further key definition:

Definition 1.3.3 *As before, suppose that $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ are two posets, and let $f \colon P \to Q$ be a map between their carrier sets. Then f is an* order-isomorphism *iff f is a surjection (is onto) and is an order-embedding.*

Note that, given our last theorem, an order-isomorphism f between $\mathscr{P}$ and $\mathscr{Q}$ is one-one and onto. Hence, 'isomorphism' is indeed an apt label. An order-isomorphism f makes one poset a copy of the other, as far as its order-structure is concerned.

And now we can sharpen our claim that, if f is an order-embedding between $\mathscr{P}$ and $\mathscr{Q}$, then f maps $\mathscr{P}$ to a copy embedded inside $\mathscr{Q}$. Stating the result with a slight abuse of notation, we have:

Theorem 1.3.4 *If $f\colon P \to Q$ is an order-embedding between $\mathscr{P} = \langle P, \preccurlyeq\rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq\rangle$, then f is an order-isomorphism between $\mathscr{P}$ and the poset $\langle f(P), \sqsubseteq\rangle$.*

$f(P)$ is, of course, the image of P under f, i.e. $\{q \mid (\exists p \in P) f(p) = q\}$. And the (very) slight abuse of notation comes in not explicitly marking that the order-relation in $\langle f(P), \sqsubseteq\rangle$ is the strictly speaking the relation in $\langle Q, \sqsubseteq\rangle$ *restricted to* $f(P)$. But here and henceforth we'll allow ourselves that degree of notational slack. The proof of our theorem is trivial: for f is, of course, onto $f(P)$, and is still an order-embedding.

Let's use our definitions, then, in a few examples:

(1) Take the pair of posets $\langle \mathbb{R}^+, \leqslant\rangle$, $\langle \mathbb{N}, \leqslant\rangle$, where $\mathbb{R}^+$ is the set of positive reals, and the order relations are the natural ones. Let $f\colon \mathbb{R}^+ \to \mathbb{N}$ map a real r in $\mathbb{R}^+$ to the natural number corresponding to r's integral part. Then f is monotone, but not an embedding (since e.g. the reals $\sqrt{10}$ and $\sqrt{11}$ get mapped to the same natural number, i.e. 3).

(2) That first example reminds us that a monotone map typically *doesn't* preserve *all* order information. For an extreme example to hammer home the point, take $\langle \mathbb{R}^+, \leqslant\rangle$ again and the one-element poset $\langle \{0\}, =\rangle$ (check that that *is* a poset!). Then the function f that maps every element of $\mathbb{R}^+$ to 0 is monotone, but it's a forgetful map that throws away *all* non-trivial order information.

(3) Take the posets $\langle \mathbb{R}^+, \leqslant\rangle$, $\langle \mathbb{N} \leqslant\rangle$ again, but this time consider a map $g\colon \mathbb{N} \to \mathbb{R}^+$ going in the opposite direction which takes a natural number to the corresponding integral real. Then g is an order-embedding but not a full isomorphism: g just finds a copy of the naturals embedded inside $\mathbb{R}^+$.

(4) Consider the poset $\langle \mathbb{R}, \leqslant\rangle$, and its order-dual $\langle \mathbb{R}, \geqslant\rangle$. Then the map $h\colon \mathbb{R} \to \mathbb{R}$ defined by $h\colon x \mapsto -x$ is an order-isomorphism.

(5) But don't be misled by that last example – there isn't always an order-isomorphism between a poset and its order-dual. Consider, for example, $\langle \mathbb{N}, \leqslant\rangle$ and *its* order-dual $\langle \mathbb{N}, \geqslant\rangle$. An order-isomorphism would have to map 0 which comes first in the $\leqslant$ order to some element which similarly comes first in the $\geqslant$ order.

But there is no such element in $\mathbb{N}$. So there can be no order-isomorphism between $\langle\mathbb{N}, \leqslant\rangle$ and $\langle\mathbb{N}, \geqslant\rangle$.

1.4 Compounding maps

Another key result is that monotone maps can be composed together to get another monotone map:

Theorem 1.4.1 *If $f\colon P \to Q$ is a monotone map from $\mathscr{P} = \langle P, \preccurlyeq\rangle$ to $\mathscr{Q} = \langle Q, \sqsubseteq\rangle$ and $g\colon Q \to R$ is a monotone map from $\mathscr{Q}$ to $\mathscr{R} = \langle R, \leqslant\rangle$, then 'g following f', i.e. the composite map $g \circ f\colon P \to R$, is a monotone map from $\mathscr{P}$ to $\mathscr{R}$.*

Proof Need we spell this out? If $p \preccurlyeq p'$ then $f(p) \sqsubseteq f(p')$, since f is monotone. And if $f(p) \sqsubseteq f(p')$ then $g(f(p)) \leqslant g(f(p'))$, since g is monotone. So if $p \preccurlyeq p'$ then $g(f(p)) \leqslant g(f(p'))$, hence $g \circ f$ is monotone. □

It is also worth remarking on the following easy result:

Theorem 1.4.2 *Composition of monotone maps is associative: in other words, if $f\colon P \to Q$, $g\colon Q \to R$, $h\colon R \to S$ are monotone, then $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are monotone and equal.*

Which leads to a brief aside. For it is, of course, trivial that the identity map $Id_P\colon P \to P$, which maps every $p \in P$ to itself, is a monotone map from $\mathscr{P} = \langle P, \preccurlyeq\rangle$ to itself. Hence – if you happen to know what a 'category' is – you'll see from our last two theorems that it's immediate that, if we take posets as objects and monotone mappings between them as the associated arrows/morphisms, then we get a category **Pos** of posets. However, although we'll sometimes be in the close neighbourhood of some category-theoretic ideas, e.g. in the next chapter, we'll not explicitly discuss categories in this book. End of aside!

To continue: Theorem 1.4.1 tells us that monotone maps compose. But so do order-embeddings and (crucially) isomorphisms:

Theorem 1.4.3 *If $f\colon P \to Q$ is an order-embedding (isomorphism) from $\mathscr{P}$ to $\mathscr{Q}$, and $g\colon Q \to R$ is an order-embedding (isomorphism) from $\mathscr{Q}$ to $\mathscr{R}$, then $g \circ f\colon P \to R$, is an order-embedding (isomorphism) from $\mathscr{P}$ to $\mathscr{R}$.*

Proof To show order-embeddings compose, just change the 'if's to 'if and only if's in the proof of Theorem 1.4.1. And to show order-isomophisms compose, we just need to remark that if f and g are onto, so is $g \circ f$. $\square$

1.5 Order similarity

To repeat, an order-isomorphism between posets is indeed a one-to-one correspondence between their carrier sets, one that preserves all the basic facts about the way that the elements are ordered. So let's say that

Definition 1.5.1 *The posets $\mathscr{P}$ and $\mathscr{Q}$ are* order-similar *(in symbols, $\mathscr{P} \cong \mathscr{Q}$), iff there is an order-isomorphism from $\mathscr{P}$ to $\mathscr{Q}$.*

And we can readily check that 'order-similarity' is properly so-called:

Theorem 1.5.2 *Order-similarity is an equivalence relation.*

Proof It is trivial that order-similarity is reflexive. Theorem 1.4.3 tells us that order-isomorphisms compose, and hence that order-similarity is transitive. So we just need to check for symmetry. In other words, we just need to confirm that if $f\colon P \to Q$ between $\mathscr{P}$ and $\mathscr{Q}$, then there is an order-isomorphism between $\mathscr{Q}$ and $\mathscr{P}$. To show that, (i) define $f^{-1}\colon Q \to P$ to be such that $f^{-1}(q) = p$ iff $f(p) = q$, and then (ii) check this is indeed an order-isomorphism between $\mathscr{Q}$ and $\mathscr{P}$. $\square$

Let's have just a very small handful of examples:

(1) We have already seen that $\langle \mathbb{N}, \leqslant \rangle$ and $\langle \mathbb{N}, \geqslant \rangle$ are of different order-types. Likewise there can be no order-isomorphism between $\langle \mathbb{N}, \leqslant \rangle$ and $\langle \mathbb{N}, \preccurlyeq \rangle$, where $\preccurlyeq$ is the evens-before-odds ordering. For in $\langle \mathbb{N}, \preccurlyeq \rangle$, 1 has an infinite number of predecessors, and there is no element in $\langle \mathbb{N}, \leqslant \rangle$ with the same order property. Hence $\langle \mathbb{N}, \leqslant \rangle$, $\langle \mathbb{N}, \geqslant \rangle$, $\langle \mathbb{N}, \preccurlyeq \rangle$ and indeed $\langle \mathbb{N}, = \rangle$ are posets with the same carrier set but of different order-types. However, $\langle \mathbb{N}, \leqslant \rangle$, $\langle \mathbb{N} \setminus \{0\}, \leqslant \rangle$ and $\langle \mathbb{E}, \leqslant \rangle$ are all order-similar (where $\mathbb{N} \setminus \{0\}$ is the set of naturals with 0 deleted, and $\mathbb{E}$ is the set of even numbers).

(2) If $\mathbb{Q}^+$ is the set of positive rational numbers, and $\leqslant$ their natural ordering, then $\langle \mathbb{Q}^+, \leqslant \rangle$ is plainly not order-similar to $\langle \mathbb{N}, \leqslant \rangle$. For, in their natural ordering, the rationals are *dense* – between any two different ones there is another – and not so for the natural

numbers. However we can also put a different order on the positive rationals. Map each $q \in \mathbb{Q}^+ \setminus \{0\}$ one-to-one to the pair of naturals $\langle m, n \rangle$ where $q = m/n$ and the fraction is in lowest terms. Do a standard kind of 'zig-zag' enumeration of such pairs $\langle m, n \rangle$ where m, n have no common divisors other than 1. In other words, say $\langle m, n \rangle$ precedes $\langle m', n' \rangle$ if $m+n \leqslant m'+n'$ or else if $m+n = m'+n'$ and $m \leqslant m'$. And now put $q \preccurlyeq q'$ in $\mathbb{Q}^+$ when either $q = 0$ or the counterpart pair for q is no later than the counterpart pair for q' in the zig-zag enumeration. Then $\langle \mathbb{Q}^+, \preccurlyeq \rangle \cong \langle \mathbb{N}, \leqslant \rangle$.

(3) The examples just mentioned involve posets whose ordering relation $\preccurlyeq$ is total (i.e., for every x, y in the carrier set, either $x \preccurlyeq y$ or $y \preccurlyeq x$). Plainly, no poset which *isn't* totally ordered can be order-similar to one that *is*.

For an example of a pair of posets which are not totally ordered but which are still order-similar, consider the following trivial example. (i) The inclusion poset $\langle \mathcal{P}(\{0, 1, 2\}), \subseteq \rangle$. (ii) The poset whose carrier set is $\{1, 2, 3, 5, 6, 10, 15, 30\}$, with the numbers ordered by divisibility. It can again be left as an exercise to specify an order-isomorphism from (i) to (ii) – there is more than one!

1.6 Inclusion posets as typical

Posets, as we said before, are ubiquitous. Evidently they can come in all manner of sizes and shapes. And they can be built up from all manner of elements, related in a huge variety of ways. It might seem quite hopeless to try to put any organization at all on the mess here.

But we can in fact tidy things up just a bit. For remember, any collection of subsets of some set A, ordered by $\subseteq$, is a poset. We dubbed this sort of poset an *inclusion poset*. And we can now show:

Theorem 1.6.1 *Every poset is order-isomorphic to an inclusion poset.*

NB, an inclusion poset can be based on *any* collection of subsets of a set A. The case where the poset is based on the set of *all* subsets of some A is a very special one. The theorem does *not* say that every poset is similar to one of the very special cases!

Proof Take the poset $\mathscr{P} = \langle P, \preccurlyeq \rangle$. For each $p \in P$, now form the set $\pi_p = \{x \mid x \in P \wedge x \preccurlyeq p\}$. Let Π be the set of all π_p for $p \in P$. Then Π is a set of subsets of P, and $\mathscr{P}^\pi = \langle \Pi, \subseteq \rangle$ is of course an inclusion poset.

Further, $\mathscr{P}$ is order-isomorphic to $\mathscr{P}^\pi$. For consider the function $f\colon P \to \Pi$ which maps $p \in P$ to $\pi_p \in \Pi$. This map is evidently a one-one correspondence between P and Π. Moreover, by definition $p \preccurlyeq p'$ iff $\pi_p \subseteq \pi_{p'}$. So we are done. $\square$

Hence, if what we really care about is just the order-type of various posets – i.e. we don't care about what sort of elements the carrier-set has, and are just highlighting the kind of ordering that is being put on a set of elements – then, without loss of generality, we can concentrate on inclusion posets as typical.

2
Galois connections

Even inclusion posets come in all sizes and shapes. Which means that there isn't a *great* deal more of interest that can be said about posets in general. Hence we are soon going to move on to talk about ordered sets which have rather more additional structure.

Still, before we do so, we need to explore another common relation that can obtain between posets, one which is weaker than order-isomorphism but which is still surprisingly rich in implications. The ideas here are particularly cute, and (among other things) they give us a lovely illustration of the way that modern mathematics likes to embed familiar facts into much more abstract general settings.

2.1 Galois connections defined

Definition 2.1.1 *Let $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ be posets. And suppose $f_*\colon P \to Q$ and $f^*\colon Q \to P$ are a pair of functions such that for all $p \in P$ and all $q \in Q$,*

$$(G) \qquad f_*(p) \sqsubseteq q \ \text{ iff } \ p \preccurlyeq f^*(q).$$

Then the pair $\langle f_*, f^* \rangle$ form a Galois connection *between $\mathscr{P}$ and $\mathscr{Q}$.*

If $\langle f_*, f^* \rangle$ is such a connection, f_* is said to be the *left adjoint* of the corresponding f^*, and f^* is the *right adjoint* of f_*.†

† Think: f_* appears to the left of its order sign in (G), and f^* to the right of *its* order sign. Alternatively, the terminology 'lower adjoint' vs 'upper adjoint' is used. Think: f_* appears on the lower side of its ordering sign, and f^* on the upper side. (I have the star lower for the lower adjoint. Annoyingly, some authors use f_* and f^* the other way about.)

Talk of adjoints is carried over from category theory. Category theorists in turn

There are plenty of serious mathematical examples (e.g. from number theory, abstract algebra and topology) of two posets with a Galois connection between them. But we really don't want to get bogged down in unnecessary mathematics at this stage; so for the moment let's just give some relatively simple cases, including a couple of logical ones.

(1) Suppose f is an order-isomorphism between $\mathscr{P}$ and $\mathscr{Q}$, so the inverse function f^{-1} is one too. Then $\langle f, f^{-1}\rangle$ is a Galois connection. For, trivially,

$$f(p) \sqsubseteq q \quad \text{iff} \quad f^{-1}(f(p)) \preccurlyeq f^{-1}(q) \quad \text{iff} \quad p \preccurlyeq f^{-1}(q).$$

(2) Let $\mathscr{P} = \langle \mathbb{N}, \leqslant\rangle$ and $\mathscr{Q}$ be $\langle \mathbb{Q}^+, \leqslant\rangle$, where in each case the order relation is standard. Put $f_*\colon \mathbb{N} \to \mathbb{Q}^+$ to be the standard embedding of the natural numbers into the rationals; and let $f^*\colon \mathbb{Q}^+ \to \mathbb{N}$ map a positive rational q to the natural corresponding to its integral part. Then (G) trivially holds, and $\langle f_*, f^*\rangle$ is a Galois connection between the integers and the rationals in their natural orderings.

(3) Take an arbitrary poset $\mathscr{P} = \langle P, \preccurlyeq\rangle$, and put $\mathscr{I} = \langle \{0\}, =\rangle$ (so $\mathscr{I}$ is a singleton poset with the only possible order relation). Let f_* be the trivial function i that maps each and every member of P to 0. And let f^* be some function that maps 0 to a particular member of P. Then f^* is right adjoint to f_*, i.e. the pair $\langle f_*, f^*\rangle$ form a Galois connection between $\mathscr{P}$ and $\mathscr{I}$, just if, for every $p \in P$, $f_*(p) = 0$ iff $p \preccurlyeq f^*(0)$, i.e. just if for every p, $p \preccurlyeq f^*(0)$, i.e. just if $f^*(0)$ is the *maximum* of P in the obvious sense. Dually, if we put f^* to be the same flattening function i, then f_* is left adjoint to it just if $f_*(0)$ is the *minimum* of P.

(4) Our next example is from elementary logic. Recall example 11 from Section 1.1. Let L be a formal language equipped with a proof-system (it could be classical or intuitionistic, or indeed any other logic with well-behaved conjunction and conditional connectives). Then $|\alpha|$ is the class of L-wffs interderivable with α; E

seem to have borrowed the term from the old theory of Hermitian operators, where in e.g. a Hilbert space with inner product $\langle\cdot,\cdot\rangle$, the operators A and A^* are said to be adjoint when we have, generally, $\langle Ax, y\rangle = \langle x, A^*y\rangle$. The formal analogy is evident.

The first discussion of such a connection – and hence the name – is to be found in Evariste Galois's work in what has come to be known as Galois theory, a topic far beyond our purview here. Though you might like to find out about Galois's short life from the historical sketch in Stewart's *Galois Theory* (1989).

is the set of all such equivalence classes; and $|\alpha| \Rightarrow |\beta|$ iff $\alpha \vdash \beta$. $\mathscr{E} = \langle E, \Rightarrow \rangle$ is a poset.

Now consider the following two functions between $\mathscr{E}$ and itself. Fix γ to be some L-wff. Then let f_* be the function that maps the equivalence class $|\alpha|$ to the class $|(\gamma \wedge \alpha)|$, and let f^* be the function that maps the equivalence class $|\alpha|$ to $|(\gamma \supset \alpha)|$. Now, we have $(\gamma \wedge \alpha) \vdash \beta$ if and only if $\alpha \vdash (\gamma \supset \beta)$. So $|(\gamma \wedge \alpha)| \Rightarrow |\beta|$ if and only if $|\alpha| \Rightarrow |(\gamma \supset \beta)|$. So $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{E}$ and itself. ('Conjunction is left adjoint to conditionalization.')

(5) Our last example for the moment is again from elementary logic. Let L now be a first-order language, and $Form(\vec{x})$ be the set of L-wffs with at most the variables $\vec{x}$ free. We'll write $\varphi(\vec{x})$ for a formula in $Form(\vec{x})$, $|\varphi(\vec{x})|$ for the class of formulae interderivable with $\varphi(\vec{x})$, and $E_{\vec{x}}$ for the set of such equivalence classes of formulae from $Form(\vec{x})$. With $\Rightarrow$ as in the last example, $\langle E_{\vec{x}}, \Rightarrow \rangle$ is a poset for any $\vec{x}$.

Now we'll consider two maps between the posets $\langle E_{\vec{x}}, \Rightarrow \rangle$ and $\langle E_{\vec{x},y}, \Rightarrow \rangle$. In other words, we are going to be moving between (classes of) formulae with at most $\vec{x}$ free, and formulae with at most $\vec{x}, y$ free (where y isn't among the $\vec{x}$).

First, since every wff with at most the variables $\vec{x}$ free also has at most the variables $\vec{x}, y$ free, there is a trivial map f_* that sends $|\varphi(\vec{x})| \in E_{\vec{x}}$ to the *same* element $|\varphi(\vec{x})| \in E_{\vec{x},y}$.

Second, we define the companion map f^* that sends $|\varphi(\vec{x}, y)| \in E_{\vec{x},y}$ to $|\forall y \varphi(\vec{x}, y)| \in E_{\vec{x}}$.

Then $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{E}_{\vec{x}}$ and $\mathscr{E}_{\vec{x},y}$. That is to say

$$f_*(|\varphi(\vec{x})|) \Rightarrow |\psi(\vec{x}, y)| \quad \text{iff} \quad |\varphi(\vec{x})| \Rightarrow f^*(|\psi(\vec{x}, y)|).$$

For that just reflects the familiar logical rule that

$$\varphi(\vec{x}) \vdash \psi(\vec{x}, y) \quad \text{iff} \quad \varphi(\vec{x}) \vdash \forall y \psi(\vec{x}, y)$$

so long as y is not free in $\varphi(\vec{x})$.

Hence universal quantification is right-adjoint to a certain trivial operation.

And we can similarly show that existential quantification is left-adjoint to the same operation (exercise: spell out that last claim).

Our first family of cases, then, shows that Galois connections exist and are at least as plentiful as order-isomorphisms. The second and third cases show that posets that aren't order-isomorphic can in fact still be connected. The third and fifth cases gives a couple of illustrations of how a more contentful notion (taking maxima, universal quantification) can be regarded as adjoint to some relatively trivial operation. The fourth case shows that even when the Galois-connected posets *are* isomorphic (in this case trivially so, because they are identical!), there can be functions f_* and f^* which aren't isomorphisms but which also go to make up a connection between the posets – and since γ was arbitrary, this case also shows that there can be many different connections between the same posets. And the fourth and fifth cases already show that Galois connections are of interest to logicians.

Finally in this section, let's show that connections are well-behaved, in the sense that connections compose to give new connections. In other words, we have:

Theorem 2.1.2 *Let $\mathscr{P} = \langle P, \preccurlyeq \rangle$, $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$, and $\mathscr{R} = \langle R, \subseteq \rangle$ be posets. Let suppose $\langle f_*, f^* \rangle$ is a Galois connections between $\mathscr{P}$ and $\mathscr{Q}$, and $\langle g_*, g^* \rangle$ is a Galois connections between $\mathscr{Q}$ and $\mathscr{R}$. Then $\langle g_* \circ f_*, f^* \circ g^* \rangle$ is a Galois connection between $\mathscr{P}$ and $\mathscr{R}$.*

Proof By assumption, we have (G_1) $f_*(p) \sqsubseteq q$ iff $p \preccurlyeq f^*(q)$ and (G_2) $g_*(q) \subseteq r$ iff $q \sqsubseteq g^*(r)$. So (G_2) gives us $g_*(f_*(p)) \subseteq r$ iff $f_*(p) \sqsubseteq g^*(r)$. And ($G_1$) gives us $f_*(p) \sqsubseteq g^*(r)$ iff $p \preccurlyeq f^*(g^*(r))$. Whence, as we want, $g_*(f_*(p)) \subseteq r$ iff $p \preccurlyeq f^*(g^*(r))$. □

2.2 An alternative definition

So far, so good. We'll meet a rather more telling example of a logically significant Galois connection in the next chapter. But first let's explore the general case in a bit more detail. Here's a simple but illuminating result:

Theorem 2.2.1 *Suppose $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ are posets, and $f_* \colon P \to Q$ and $f^* \colon Q \to P$ are a pair of functions between their carrier sets. Then $\langle f_*, f^* \rangle$ is a Galois connection if and only if*

(i) *f_*, f^* are both monotone, and*
(ii) *for all $p \in P$, $q \in Q$, $p \preccurlyeq f^*(f_*(p))$ and $f_*(f^*(q)) \sqsubseteq q$.*

Proof (*Only if*) Suppose $\langle f_*, f^*\rangle$ is a Galois connection. As a particular case of (G) from Defn. 2.1.1 we have $f_*(p) \sqsubseteq f_*(p)$ iff $p \preccurlyeq f^*(f_*(p))$. Since $\sqsubseteq$ is reflexive, the l.h.s. of that biconditional holds. So $p \preccurlyeq f^*(f_*(p))$. Similarly for the other half of (ii).

Now, suppose also that $p \preccurlyeq p'$. Since we've shown $p' \preccurlyeq f^*(f_*(p'))$, we have $p \preccurlyeq f^*(f_*(p'))$. But by (G) we have $f_*(p) \sqsubseteq f_*(p')$ iff $p \preccurlyeq f^*(f_*(p'))$. Whence, $f_*(p) \sqsubseteq f_*(p')$ and f_* is therefore monotone. Similarly for the other half of (i).

(*If*) Now assume (i) and (ii) hold, and suppose $f_*(p) \sqsubseteq q$. Since by (i) f^* is monotone, $f^*(f_*(p)) \preccurlyeq f^*(q)$. But by (ii) $p \preccurlyeq f^*(f_*(p))$. So, since $\preccurlyeq$ is transitive, $p \preccurlyeq f^*(q)$. Which establishes that if $f_*(p) \sqsubseteq q$ then $p \preccurlyeq f^*(q)$. The proof of the other half of the biconditional (G) is dual. □

This theorem can evidently be used to justify an alternative definition of a Galois connection as, simply, *a pair of maps for which conditions (i) and (ii) hold.*‡

2.3 The relation between adjoints

Our next theorem tells us that the relation between the adjoint members of a Galois connection is 'rigid' in the sense that if $\langle f_*, f^*\rangle$ is to be a connection, then f_* fixes what f^* uniquely has to be, and conversely f^* fixes what f_* has to be.

Theorem 2.3.1 *If $\langle f_*, f^1\rangle$ and $\langle f_*, f^2\rangle$ are Galois connections between $\langle P, \preccurlyeq\rangle$ and $\langle Q, \sqsubseteq\rangle$, then $f^1 = f^2$. Likewise, if $\langle f_1, f^*\rangle$ and $\langle f_2, f^*\rangle$ are Galois connections between the same posets, then $f_1 = f_2$.*

‡ Because (i) holds, connections defined our way are sometimes called *monotone* Galois connections.

But now recall that if $\mathscr{Q} = \langle Q, \sqsubseteq\rangle$ is a poset, so is its dual $\mathscr{Q}^{op} = \langle Q, \sqsubseteq^{op}\rangle$. Suppose then that $\langle f_*, f^*\rangle$ is a (monotone) connection between $\mathscr{P} = \langle P, \preccurlyeq\rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq\rangle$ so that $f_*(p) \sqsubseteq q$ iff $p \preccurlyeq f^*(q)$. Then, trivially, there is an *antitone* Galois connection between $\mathscr{P}$ and $\mathscr{Q}^{op}$, meaning a pair of functions $\langle f_*, f^*\rangle$ such that $q \sqsubseteq^{op} f_*(p)$ iff $p \preccurlyeq f^*(q)$.

The only point of footnoting this here is to remark that, (a) you might well encounter the idea of a connection defined the antitone way (which indeed it how it appears in Galois's work); indeed (b) some authors only call the antitone case a Galois connection, reserving the term adjunction for connections defined our way. But (c) since a monotone connection between $\mathscr{P}$ and $\mathscr{Q}$ is an antitone connection between $\mathscr{P}$ and $\mathscr{Q}$'s dual, we just don't need to fuss about the difference, and without loss of generality can we can concentrate entirely on connections presented the monotone way.

Proof From (G) applied to the connection $\langle f_*, f^1 \rangle$, putting $p = f^2(q)$, we get $f_*(f^2(q)) \sqsubseteq q$ iff $f^2(q) \preccurlyeq f^1(q)$. But by Theorem 2.2.1 applied to the connection $\langle f_*, f^2 \rangle$, we have $f_*(f^2(q)) \sqsubseteq q$. Hence $f^2(q) \preccurlyeq f^1(q)$. Similarly from (G) applied to $\langle f_*, f^2 \rangle$ and Theorem 2.2.1 applied to $\langle f_*, f^1 \rangle$ we get $f^1(q) \preccurlyeq f^2(q)$. Hence $f^1(q) = f^2(q)$. But q was arbitrary in Q. So $f^1 = f^2$. Dually, of course, for the other part of the theorem. □

Careful, though! This theorem does *not* say that, for any f_* which maps between the carrier sets of $\mathscr{P}$ and $\mathscr{Q}$, there must actually exist a unique corresponding f^* such that $\langle f_*, f^* \rangle$ form a Galois connection (we'll note counterexamples in a moment). Nor does it say that when there *is* a Galois connection between two given posets $\mathscr{P}$ and $\mathscr{Q}$, it is unique (our toy examples in Sec. 2.1 already showed that *that* is false). The claim is only that, if you are given a possible left adjoint (or a possible right adjoint), there can be at most one candidate for its companion to complete a connection.

Given that adjoint functions determine each other, we naturally seek an explicit definition of one in terms of the other. Here it is:

Theorem 2.3.2 *If $\langle f_*, f^* \rangle$ is a Galois connection between $\langle P, \preccurlyeq \rangle$ and $\langle Q, \sqsubseteq \rangle$, then*

(i) $f^*(q) =$ *the maximum of* $\{p \in P \mid f_*(p) \sqsubseteq q\}$
(ii) $f_*(p) =$ *the minimum of* $\{q \in Q \mid p \preccurlyeq f^*(q)\}$

Proof Here, 'maximum' and 'minimum' are, again, the obvious notions. So, m is a maximum of the poset $\langle P, \preccurlyeq \rangle$ if $m \in P$, and for all $p \in P$, $p \preccurlyeq m$. And maxima are unique when they exist (for if m and m' are both maxima of P, $m' \preccurlyeq m$ and $m \preccurlyeq m'$ so $m = m'$). Dually for minima.

To show (1), first recall Theorem 2.2.1 shows that $f_*(f^*(q)) \sqsubseteq q$. Therefore $f^*(q) \in \{p \in P \mid f_*(p) \sqsubseteq q\}$. Now suppose p is such that $f_*(p) \sqsubseteq q$. Since f^* is monotone, it follows that $f^*(f_*(p)) \preccurlyeq f^*(q)$. But by Theorem 2.2.1 again, $p \preccurlyeq f^*(f_*(p))$, so $p \preccurlyeq f^*(q)$. Hence $f^*(q)$ has to be the maximum member of $\{p \in P \mid f_*(p) \sqsubseteq q\}$.

The proof of (2) is dual. □

It follows that if f_* is a function such that the set $\{p \in P \mid f_*(p) \sqsubseteq q\}$ *doesn't* always have a maximum, then there can't be a corresponding right adjoint f^*. For example, put $\mathscr{P} = \langle \mathbb{N}, \preccurlyeq \rangle$, let $\mathscr{I} = \langle \{0\}, = \rangle$ be a singleton poset, and suppose f_* is the trivial function that maps each member of P to $*$. Then f_* has no right adjoint (see Example (iii) in the last section). For another example, put $P = \mathbb{Q}^+$, and $Q = \mathbb{N}$ and give each carrier set its natural order. For $p \in \mathbb{Q}^+$, put $f_*(p)$ to be the natural corresponding to p's integral part. Then, e.g., for every rational p such $1 \leqslant p < 2$, $f_*(p) = 1$ and hence $f_*(p) \sqsubseteq 1$, and there is no maximum member of $\{p \in \mathbb{Q}^+ \mid f_*(p) \sqsubseteq 1\}$ since there is no maximum rational less than 2. Hence there is no right adjoint to this function.

It also immediately follows, by the way, that

Theorem 2.3.3 *Galois connections are not necessarily symmetric. That is to say, given $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{P}$ and $\mathscr{Q}$, it does* not *follow that $\langle f^*, f_* \rangle$ is a connection between $\mathscr{Q}$ and $\mathscr{P}$.*

Proof Use Example 2 in Sec. 2.1; take the Galois connection defined there between $\mathscr{P} = \langle \mathbb{N}, \leqslant \rangle$ and $\mathscr{Q}$ be $\langle \mathbb{Q}^+, \leqslant \rangle$, with the right adjoint mapping a rational to the natural corresponding to its integral part. We have just proved, however, that that latter function can't also appear as a left adjoint in a connection from $\mathscr{Q}$ to $\mathscr{P}$. □

2.4 Fixed points and closures

Recall again that we use $f(P)$ for the image of the set P under f, i.e. $f(P) = \{f(p) \mid p \in P\}$. And p is a fixed point of a function f if $f(p) = p$. Then we have the following:

Theorem 2.4.1 *If $\langle f_*, f^* \rangle$ is a Galois connection between $\langle P, \preccurlyeq \rangle$ and $\langle Q, \sqsubseteq \rangle$, then*

(i) *$f_* \circ f^* \circ f_* = f_*$ and $f^* \circ f_* \circ f^* = f^*$,*

(ii) *$p \in f^*(Q)$ if and only if p is a fixed point of $f^* \circ f_*$; and $q \in f_*(P)$ if and only if q is a fixed point of $f_* \circ f^*$.*

(iii) *$f^*(Q) = f^*(f_*(P))$ and $f_*(P) = f_*(f^*(Q))$.*

Proof (1) We have shown that, for all $p \in P$, $p \preccurlyeq f^*(f_*(p))$. Since f_* is monotone, $f_*(p) \sqsubseteq f_*(f^*(f_*(p)))$. But also, by an instance of (G), we have $f_*(f^*(f_*(p))) \sqsubseteq f_*(p)$ iff $f^*(f_*(p)) \preccurlyeq f^*(f_*(p))$. Since the r.h.s. of that biconditional is true because $\preccurlyeq$ is reflexive, we thus also have $f_*(f^*(f_*(p))) \sqsubseteq f_*(p)$. Whence $f_*(f^*(f_*(p))) = f_*(p)$, for $\sqsubseteq$ is antisymmetric. Similarly for the other half of (1).

(2) Suppose $p \in f^*(Q)$. Then, by definition, for some q, $p = f^*(q)$. Hence $f^*(f_*(p)) = f^*(f_*(f^*(q))) = f^*(q) = p$. Which is just what it

means to say that p is a fixed point of $f^* \circ f_*$. Similarly if $q \in f_*(P)$, then it is a fixed point of $f_* \circ f^*$.

Suppose conversely that p is a fixed point of $f^* \circ f_*$, so $p = f^*(f_*(p))$. Then p is the value of $f^*(q)$ where $q = f_*(p) \in Q$, so $p \in f^*(Q)$. Similarly, if q is a fixed point of $f_* \circ f^*$, then $q \in f_*(P)$.

(3) We've just shown that if $p \in f^*(Q)$, then $p = f^*(f_*(p))$, so $p \in f^*(f_*(P))$. Hence $f^*(Q) \subseteq f^*(f_*(P)$. Conversely, suppose $p \in f^*(f_*(P))$: then $p = f^*(q)$ for some $q \in f_*(P) \subseteq Q$. So $p \in f^*(Q)$, and $f^*(f_*(P)) \subseteq f^*(Q)$. Hence $f^*(f_*(P)) = f^*(Q)$.

Similarly for the other half of (3). □

Now let's introduce a new bit of abbreviatory notation:

Definition 2.4.2 *Given a Galois connection between $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \subseteq \rangle$, let $\mathscr{P}^f$ be $\mathscr{P}$'s sub-poset $\langle f^*(f_*(P)), \preccurlyeq \rangle$, where $\preccurlyeq$ here is $\mathscr{P}$'s order relation restricted to $f^*(f_*(P))$. Similarly, put $\mathscr{Q}_f$ for the corresponding sub-poset $\langle f_*(f^*(Q)), \sqsubseteq \rangle$.*

Then our last theorem can be used to prove a more consequential one:

Theorem 2.4.3 *If $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{P} = \langle P, \preccurlyeq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$, then $\mathscr{P}^f$ and $\mathscr{Q}_f$ are order-isomorphic.*

We know that a pair of posets which have a Galois connection between them needn't be isomorphic overall, and can – so to speak – be lopsidedly related. But this nice theorem says that they must, for all that, contain a pair of isomorphic sub-posets. Moreover, as we will see, we can extract the isomorphic map between those contained sub-posets from the connection.

Proof By Theorem 2.4.1, $\mathscr{P}^f = \langle f^*(Q), \preccurlyeq \rangle$ and $\mathscr{Q}_f = \langle f_*(P), \preccurlyeq \rangle$. We'll show that f_* restricted to $f^*(Q)$ provides the required order-isomorphism from $\mathscr{P}^f$ to $\mathscr{Q}_f$.

For a start, note that since $f^*(Q) \subseteq P$, f_* restricted to $f^*(Q)$ takes elements into $f_*(P)$. So f_* is indeed a map between the carrier sets of $\mathscr{P}^f$ and $\mathscr{Q}_f$

We now need to check (i) that $f_* \colon f^*(Q) \to f_*(P)$ is onto and (ii) it is an order-embedding.

(i) We've just shown that $f_* \circ f^* \circ f_* = f_*$. So every element in $f_*(P)$, i.e. every element $f_*(p)$ for $p \in P$, is *also* the value of f_* for the argument $f^*(f_*(p))$, which is an argument in $f^*(Q)$. So f_* is onto $f_*(P)$.

(ii) To say $f_*\colon f^*(Q) \to f_*(P)$ is an order-embedding is to say that for all $p, p' \in f^*(Q)$, if $p \preccurlyeq p'$ then $f_*(p) \sqsubseteq f_*(p')$ *and* if $f_*(p) \sqsubseteq f_*(p')$ then $p \preccurlyeq p'$. The first conjunct is immediate, as we know that the original $f_*\colon P \to Q$ is monotone, so its restriction to $f^*(Q)$ is monotone too. For the second conjunct, suppose $p, p' \in f^*(Q)$ and $f_*(p) \sqsubseteq f_*(p')$. Then, since f^* is monotone, $f^*(f_*(p)) \preccurlyeq f^*(f_*(p'))$. But we've just proved that $f^* \circ f_*$ maps an element to itself inside $f^*(Q)$. So, for $p, p' \in f^*(Q)$, $p \preccurlyeq p'$. And we are done. □

Definition 2.4.4 *Suppose* $\mathscr{P} = \langle P, \preccurlyeq\rangle$ *is a poset; then a* closure function *for* $\mathscr{P}$ *is a function* c *such that, for all* $p, p' \in P$,

(i) $p \preccurlyeq c(p)$;
(ii) *if* $p \preccurlyeq p'$, *then* $c(p) \preccurlyeq c(p')$, *i.e.* c *is monotone;*
(iii) $c(c(p)) = c(p)$ *i.e.* c *is 'idempotent'.*

Roughly speaking, then, a closure function c maps a poset 'upwards' to a subset which then stays fixed under further applications of c. It is easy to show:

Theorem 2.4.5 *If* $\langle f_*, f^*\rangle$ *is a Galois connection between* $\mathscr{P} = \langle P, \preccurlyeq\rangle$ *and some poset, then* $f^* \circ f_*$ *is a closure function for* $\mathscr{P}$.

Proof We quickly check that the three conditions for closure apply. We have (1) $p \preccurlyeq f^*(f_*(p))$ by Theorem 2.2.1.

(2) $p \preccurlyeq p'$ implies $f_*(p) \sqsubseteq f_*(p')$ (where $\sqsubseteq$ is the ordering of the connected poset) since f_* is monotone. Which implies that $f^*(f_*(p)) \preccurlyeq f^*(f_*(p'))$, since f^* is monotone.

(3) Theorem 2.4.1 tells us that $f_* \circ f^* \circ f_* = f_*$. It immediately follows that $(f^* \circ f_*) \circ (f^* \circ f_*) = f^* \circ f_*$. □

2.5 One way a Galois connection can arise

We've so far mostly been talking about Galois connections in general. Let's finally note one particular way by which connections can arise.

Theorem 2.5.1 *Suppose* γ, δ *are sets, and let* R *be any relation between their elements. Define a function* f_R *from subsets of* γ *to subsets of* δ *as follows: if* $\alpha \subseteq \gamma$, *then* $f_R(\alpha)$ *is the set of things which are* R*-related to everything in* α. *In other words,* $f_R(\alpha) = \{b \mid \forall a(a \in \alpha \to aRb)\}$. *And* *define a corresponding function f^R back from subsets of δ to subsets of γ like this: if $\beta \subseteq \delta$ then $f^R(\beta) = \{a \mid \forall b(b \in \beta \to aRb)\}$.*

Then $\langle f_R, f^R \rangle$ is a Galois connection between the inclusion posets $\langle \mathcal{P}(\gamma), \subseteq \rangle$ and $\langle \mathcal{P}(\delta), \supseteq \rangle$.

Proof We just have to check that (G) holds, i.e. $f_R(\alpha) \supseteq \beta$ iff $\alpha \subseteq f^R(\beta)$, by unpacking definitions.

But $f_R(\alpha) \supseteq \beta$ is equivalent to $(\forall b \in \beta)(\forall a \in \alpha)aRb$, which is equivalent to $(\forall a \in \alpha)(\forall b \in \beta)aRb$, i.e. $\alpha \subseteq f^R(\beta)$. $\square$

Let's say that such a connection – between posets involving the *powersets* of some sets γ and δ, induced by a relation R between the elements of the original γ and δ – is *relation-generated*. Galois's original example was of this kind. The example we will be looking at in the next chapter is relation-generated too.

3
The Galois connection between syntax and semantics

And now, at last, we get to explicate Lawvere's talk 'the familiar Galois connection between sets of axioms and classes of models, for a fixed [signature]'.

3.1 Some reminders

First, in this section, some quick reminders. Let L be a formal language. Then a set of L-axioms in the wide sense that Lawvere is using is just *any* old set α of L-sentences. And by talk of models, Lawvere in fact just means structures apt for interpreting a language with L's signature. So the claim is that we can make a Galois connection between *sets of sentences* and *classes of structures*.

To take a familiar type of example. Suppose L is a classical first-order language. L will have its distinctive non-logical vocabulary. The particular choice of symbolism is unimportant, of course – what really matters is L's signature, which specifies L as having a certain fixed number of constants, a certain fixed number of predicates (each with a given arity), and a certain fixed number of function-symbols (again each with a given arity). A structure apt for interpreting L will have a domain, with distinguished elements assigned to each constant, sets of n-tuples from the domain assigned to n-ary predicates, and appropriate sets of $n+1$-tuples from the domain assigned to n-ary functions. We can then define what it is for an L-sentence φ to be true with respect to the L-structure s in the entirely familiar way.

But nothing that follows will in fact depend on L's being classical or being first-order. We just need there to be *some* conception of a class of L-structures apt for interpreting L-sentences, and the idea of a particular L-sentence being true with respect to a structure.

We'll use entirely familiar notation and terminology. For the record:

Definitions 3.1.1 *For a given formal language L, φ be an L-sentence, α a set of L-sentences, and s be L-structure. Then*

(i) *If φ is true w.r.t. s, we write '$s \vDash \varphi$'.*
(ii) *The L-structure s is a* model *of α iff for every $\varphi \in \alpha$, $s \vDash \varphi$.*
(iii) *If every model of α makes φ true, we write '$\alpha \vDash \varphi$'.*

NB, contrast the second definition – the modern conventional one – with Lawvere's wider usage of 'model'. The overloading of the symbol '$\vDash$' is of course entirely standard.

3.2 Making the connection

Now we make Lawvere's connection. Make the same assumption that L is a formal language, and L-structures are those which are apt for interpreting it. Then,

Definitions 3.2.1 *Let A ('A' for axioms) be the set of all L-sentences, and S be the set of all L-structures. Then put*

(i) $\mathscr{A} = \langle \mathcal{P}(A), \subseteq \rangle$.
(ii) $\mathscr{S} = \langle \mathcal{P}(S), \supseteq \rangle$.
(iii) *For $\alpha \subseteq A$, put $f_*(\alpha) = \{s \mid \forall\varphi(\varphi \in \alpha \to s \vDash \varphi)\}$.*
(iv) *For $\sigma \subseteq S$, put $f^*(\sigma) = \{\varphi \mid \forall s(s \in \sigma \to s \vDash \varphi)\}$.*

(Ok, there's a wrinkle there which might immediately strike you: but if you spot it, bear with me, and I'll return to the point.)

Here, then, f_* is the natural 'find the models' function. It takes a bunch of sentences α and looks for the biggest collection of structures that make the sentences in the bunch all true together, i.e. it returns all the models of α. In the other direction, f^* is the natural 'find all the true sentences' function. It takes a bunch of L-structures and looks for the biggest bunch of L-sentences that are true of all of those structures. And by Theorem 2.5.1, it is immediate that

Theorem 3.2.2 *$\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{A}$ and $\mathscr{S}$.*

Proof It's the connection generated by the converse of the relation $\vDash$, which holds between elements of A and S. □

3.3 Drawing the consequences

Terrific! Now we can just grind the handle, and apply all those general theorems about Galois connections from the last chapter to our special case of the connection between $\mathscr{A}$ and $\mathscr{S}$. So let's collect together some of the implications:

Theorem 3.3.1 *With f_*, f^* as defined, forming a Galois connection between $\mathscr{A} = \langle \mathcal{P}(A), \subseteq \rangle$ and $\mathscr{S} = \langle \mathcal{P}(S), \supseteq \rangle$,*

(i) *f_* is monotone, i.e. if $\alpha \subseteq \alpha'$ then $f_*(\alpha) \supseteq f_*(\alpha')$.*
(ii) *$\alpha \subseteq f^*(f_*(\alpha))$,*
(iii) *$f_* \circ f^* \circ f_* = f_*$,*
(iv) *$\alpha \in f^*(\mathcal{P}(S))$ if and only if α is a fixed point of $f^* \circ f_*$.*

And dually we have

(v) *f^* is monotone, i.e. if $\mu \supseteq \mu'$ then $f^*(\mu) \subseteq f^*(\mu')$.*
(vi) *$f_*(f^*(\mu)) \supseteq \mu$,*
(vii) *$f^* \circ f_* \circ f^* = f^*$,*
(viii) *$\mu \in f_*(\mathcal{P}(A))$ if and only if μ is a fixed point of $f_* \circ f^*$.*

The proofs of all these claims are already to hand from the last chapter.

But what do they really *mean*? Well, result (i) just reminds us that if the set of sentences α is contained in α', then the set of models of α contains those of α'. In other words, as we expand a set of sentences we can't increase the number of ways of making them all true together. (Trivial example: expand the usual set of axioms for group theory by adding the axiom that for all elements x, y, $x \cdot y = y \cdot x$, where '$\cdot$' is group multiplication. Plainly we thereby reduce the class of models to just the class of Abelian groups.)

What about results (ii) to (iv)? To get a handle on these, let's consider the significance of the composite map $f^* \circ f_*$.

By definition, $f_*(\alpha)$ is the set of all structures which are models for α. So $f^*(f_*(\alpha))$ is the most inclusive set of sentences which are true on all those structures which are models for α. In more familiar terms, $\varphi \in f^*(f_*(\alpha))$ iff every interpretation which makes all of α true makes φ true. *Hence $f^*(f_*(\alpha))$ is just the set of logical consequences of α.*

And here we link up with a very familiar old idea. For recall:

Definitions 3.3.2 *A set of sentences θ is closed under logical consequence just in case, if $\theta \vDash \varphi$, then $\varphi \in \theta$. And then*

(i) *An L-theory θ is a set of L-sentences closed under logical consequence.*
(ii) *The theory with axioms α is the smallest theory containing α.*†

We consequently have not only that $f^* \circ f_*$ is a closure function (by Theorem 2.4.5), but it generates a closure under logical consequence – more precisely, it generates the smallest closure:

Theorem 3.3.3 $f^*(f_*(\alpha))$ *is the theory with axioms α.*

Proof We've already seen why $f^*(f_*(\alpha))$ must be closed under logical consequence and hence is a theory. By Theorem 3.3.1 (ii), that theory contains α. And it is the smallest theory that contains α. In other words, if θ is any theory which contains α, then $f^*(f_*(\alpha)) \subseteq \theta$.

Why? By hypothesis $\alpha \subseteq \theta$. Since f_* is monotone by Theorem 3.3.1 (i), $f_*(\alpha) \supseteq f_*(\theta)$. Since f^* is monotone by (v), $f^*(f_*(\alpha)) \subseteq f^*(f_*(\theta))$. But θ is a theory. And for any theory, evidently $f^*(f_*(\theta))$, the set of logical consequences of θ, is just θ. So $f^*(f_*(\alpha)) \subseteq \theta$. □

And looked at through the lens of that last theorem, parts (ii) to (iv) of Theorem 3.3.1 now become near trivia. Thus (ii) just says again that, given a set of sentences as axioms, forming the theory with those axioms can't give us a smaller set; (iii) says that if you first round out a bunch of sentences α to get the theory $\theta(\alpha)$ and then look for the theory's models, you get to the same place as if you'd just looked for the models of α straight off; and (iv) tells us that the function f^* that looks for *all* the sentences made true across a bunch of models must return as its value a set of sentences closed under logical consequence.

What about the rest of Theorem 3.3.1? Well, (v) just confirms our expectations again. It says that if the set of structures μ contains μ', then the set of truths verified by all the structures in μ is contained in the set of truths verified by all the structures in μ'. But as for the other results, what's *their* significance? What does the map $f_* \circ f^*$ do for us?

Well, $f^*(\mu)$ is the most inclusive set of L-sentences made true by every structure in μ. So $f^*(\mu)$ is a theory (that because the logical consequences of any sentences in $f^*(\mu)$ will also be made true by every structure in μ, so $f^*(\mu)$ is closed under logical consequence). And hence

† 'Theories' are, of course, equally often defined as sets of sentences closed under some syntactic deducibility relation (rather than as sets closed under semantic consequence). So let's emphasize: it is the *semantic* relation that is in play in our definition here.

$f_*(f^*(\mu))$ is the set of *all* structures which are models for that theory. Which links up with another familiar idea:

Definition 3.3.4 *An class of S-structures μ is axiomatizable if there is an L-theory θ such $m \in \mu$ iff m is a model of θ.*

We then have

Theorem 3.3.5 $f_*(f^*(\mu))$ *is an axiomatizable class of structures, and it is the smallest axiomatizable class containing μ.*

Proof Dual to the proof of Theorem 3.3.3 □

And looked at through the lens of *this* theorem, the last three parts of Theorem 3.3.1 again turn into near trivia. Thus, (ii) just repeats that the smallest axiomatizable class containing μ contains μ; (iii) says that if you first round out a bunch of models μ to get the minimal axiomatizable class and then look an axiomatization for those models, you get to the same place as if you'd just looked for the theory for μ straight off; and (iv) tells us that the function f_* that looks for *all* the models of a theory must return an axiomatizable class.

3.4 Triviality?

'Hold on! Is that it? But those results really *are* just trivial! Have we laboured so hard to bring forth such a mouse?'

An understandable first reaction, perhaps, but one that rather misses the point. So two comments.

First, the appearance of mere triviality comes when we look at the elements of Theorem 3.3.1 in the light of Theorems 3.3.3 and 3.3.5. But remember that those latter linking theorems are not themselves entirely trivial.

But second, and more importantly, we are not in any case in the business here of proving exciting new results about theories: rather we are trying to fit familiar old thoughts into a less familiar but much more general order-theoretic framework. Look at it this way. Start from the *true-of* relation which can obtain between an L-sentence and an L-structure. That immediately generates a Galois connection between two naturally ordered posets whose elements are sets of sentences and sets of structures. And this *already* dictates that e.g. the composite map $f^* \circ f_*$ has to have a special significance. So in this way, the notion of *the theory* *with axioms* α, with all the properties we'd expect of that notion, is (as it were) forced upon us, generated by a construction that appears all over the place in mathematics.

This kind of setting of familiar ideas into a wider abstract framework – so we get to see local results in one domain as in fact exemplifying a very general pattern that can be found across many domains – is characteristic of modern mathematics (category theory, perhaps, giving us the most spectacular examples). And showing in this way how the local fits into a general pattern is one kind of explanatory exercise, revealing how the particular case exemplifies a 'natural kind' of phenomenon.

I suspect that if you aren't even just a little intrigued by the elegance of this kind of pattern-finding exercise, then perhaps certain types of mathematics aren't really for you!

3.5 More consequences

It will perhaps help us to see better what's going on here when we consider Theorem 2.4.3 again, which implies in the present case that

Theorem 3.5.1 $\mathscr{A}^f = \langle f^*(f_*(\mathcal{P}(A))), \subseteq\rangle$ *and* $\mathscr{S}_f = \langle f_*(f^*(\mathcal{P}(S))), \supseteq\rangle$ *are order-isomorphic, and (the restriction of)* f_* *provides an isomorphism between them.*

So $\mathscr{A}^f$ is the poset of *theories* built in the language L, ordered by inclusion. This poset evidently has a maximum, namely the inconsistent theory containing *all* L-sentences. Then, at the top of the poset in the ordering, just under the maximum, will be those theories θ such that adding even one more new axiom to θ which isn't already in the set takes us back to the maximal, inconsistent theory. Such a θ is *consistent*, since it doesn't contain all L-sentences (here we are assuming that the semantic consequence relation for L has the usual *ex false quodlibet* property). But θ is *negation-complete*, i.e. for every L-sentence φ either $\varphi \in \theta$ or $\neg\varphi \in \theta$. (Proof of negation completeness: Suppose for reductio that, for some φ, neither $\theta \nvDash \varphi$ and $\theta \nvDash \neg\varphi$. Since $\theta \nvDash \varphi$, $\varphi \notin \theta$. Since $\theta \nvDash \neg\varphi$, the new theory with axioms θ, φ is still consistent so not the maximum theory. Which contradicts the assumption that adding any new axiom to θ gives us the maximum theory.)

As we go down the poset $\mathscr{A}^f$ further in the ordering, then we move from the maximum theory, through the negation-complete theories, on to more and more partial theories, till we get down to the empty theory as the minimum.

Now, our theorem tells us that f_* is an order-isomorphism between $\mathscr{A}^f$, our poset of theories, and a corresponding poset of axiomatizable-sets-of-structures, $\mathscr{S}_f$. And how is that second poset built up?

Well, start with the maximum of $\mathscr{A}^f$, the inconsistent theory: then f_* will map that across to the maximum of $\mathscr{S}_f$, namely the empty set of structures (remember which way up the ordering is on the structures side of the Galois connection). And then, just below the maximum of $\mathscr{A}^f$, f_* maps each consistent negation-complete theory to the corresponding set of its models. But the models of a negation-complete theory θ agree on the truth-value of *every* L-sentence (for they make φ true if $\varphi \in \theta$ and φ false if $\neg\varphi \in \theta$, and by negation completeness one or other case must hold). So f_* maps the negation-complete theories to sets of *elementarily equivalent* structures, in the model-theorist's sense.

Then, as we go further down the poset $\mathscr{A}_f$ we get more and more partial theories which settle the truth-values of more and more limited classes of sentences. And f_* maps these partial theories to bunches of structures, i.e. elements of $\mathscr{S}^f$, which agree on narrower and narrower classes of sentences closed under logical consequence. These axiomatizable bunches of structures are, by our theorems, exactly the ones which can be represented as $f_*(f^*(\mu))$ for some μ or other.

3.6 Posets as sets

And I think that's probably about as much juice as we can usefully squeeze out of idea of a Galois connection, at least as applied to the simple observation that the 'true of' relation generates a connection between sets of L-sentences and sets of L-structures. But we'll see other examples of logically salient connections in what follows.

However, we ought to finish by making one general comment, if only to quiet the worry that might well have occurred to you (well, probably *should* have occurred to you) when confronted with Definitions 3.2.1.

We said at the outset that a poset is a *set* equipped with a partial ordering. However, although that was entirely standard, it was in a sense overkill. Strictly speaking, for most of our discussion we might have done better to use a plural idiom and say that our concern is with cases where we have *some objects* and a partial ordering defined over them. For we don't in general really need to think of those many objects as themselves constituting a new object, the *set* of them.

However, we *have* been helping ourselves throughout to the idiom of set-talk, even if it arguably commits us to a bit more than we often need. For the idiom is utterly familiar, and it would indeed by now seem distractingly perverse to go out of our way to avoid it. The implicit extra commitment to many-objects-as-one-set which comes with our set-talk is mostly entirely harmless.

Though not quite always. Reflect that *any* non-empty set can be made into e.g. a structure for interpreting a first-order language with a given signature: just take the set as the domain, select out enough elements (repetitions are allowed!), and define appropriate relations and functions over the set. And now think again about our definition of $\mathscr{S}$. Its carrier 'set' is supposed to be the collection of all sets of L-structures. But that could be as big as the universe of all sets, period. So the carrier 'set' for $\mathscr{S}$, on standard views, is then too big to be a kosher set. It's a '(proper) class', which is no doubt why Lawvere did talk here of '*classes* of models'. So $\mathscr{S}$ can't really be a po*set*.

Well, no matter. We could talk about 'po-classes' instead (or we could borrow John Conway's rather nice habit of using capitalizing talk of Sets when they are really too big to be sets, and similarly talk about Posets). Alternatively, if you think the very idea of proper classes is a bit of a cheat, we could avoid referring to a set of sets-of-L-structures, and speak – at the price of some contortions – using a consistently plural idiom instead. But anyway, the point remains that, even when using set talk as we have done, we don't ever need to lean very hard on the presumption that all the collections of L-structures can themselves be collected into a new object, the over-sized Set $\mathcal{P}(S)$. Talking of $\mathscr{S}$ as a po*set* as we did, at least in the present context, is just a convenience, which we can insouciantly allow ourselves.