1

Partially ordered sets

Ordinals are used to describe the length of iterated discrete processes. Or to put it another way – for this comes to much the same (as we'll see later) – they are used to describe well-ordered sets.

So, before tackling the ordinals themselves, we first need to explore the idea of a well-ordered set. And before talking about *that*, we'd better say something about ordered sets and about the relations that can obtain between ordered sets more generally. Which is what we start doing in this opening chapter.

1.1 Posets defined

We begin by defining a type of ordered set which shows up in a wide variety of contexts:

Definition 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* $x \preccurlyeq z$*.*

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 *A poset contains no finite cycles.*

By a finite cycle, we mean a set of elements $p_0, p_1, p_2, \ldots, p_n \in P$, *all different*, such that $p_0 \preccurlyeq p_1 \preccurlyeq p_2 \preccurlyeq \ldots \preccurlyeq p_n \preccurlyeq p_0$ (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 $p_0 \preccurlyeq p_n$. But we are already given $p_n \preccurlyeq p_0$. So by anti-symmetry we can conclude $p_n = p_0$, contradicting the assumption that the elements of the cycle are all distinct. □

What about infinite cycles? Well, they are ruled out too, as we'll prove later.

Suppose we draw a diagram of a poset: put a dot for each element, and a directed arrow from the dot corresponding to p to the dot corresponding to p' just when $p \preccurlyeq p'$. Then the diagram will have no loops where we can follow a linked chain of 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 are just a few examples of posets, starting with some based on familiar number systems, and then broadening out:

(1) The natural numbers $\mathbb{N}$ with $\preccurlyeq$ read as the '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 a partial ordering (the 'smallest' partial ordering on any set).
(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) The set of all 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) The powerset $\mathcal{P}(A)$ of some set A, with the members of $\mathcal{P}(A)$ ordered by set-inclusion $\subseteq$.

(12) Choose a propositional 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.

(13) A downward-branching binary tree has a top node, and every node has zero, one or two children (logicians: 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.

(14) Just for 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.

Note, by the way, that we can generalize our example 10. Set-inclusion is of course always reflexive, anti-symmetric and transitive, therefore we have:

(15) *Any* arbitrary set $\mathcal{P}(M)(A)$ of subsets of a set A, ordered by $\subseteq$ again, is a poset.

Call such cases *inclusion posets.* A bit later, Theorem 1.10 will tell us that such inclusion posets are in a good sense typical.

Note also that a poset's ordering relation $\preccurlyeq$ may in fact be *total*, i.e. it may be the case that for every x, y in the carrier set, either $x \preccurlyeq y$ or $y \preccurlyeq x$ (as in case 1, though not in e.g. cases 5 and 10). Perhaps we should really call our orders 'at least partial': but plain 'partial' is absolutely standard.

Another elementary point is illustrated by the relationship between examples 5 and 6, and by the relationship between examples 11 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 when the carrier set is $\mathcal{P}(A)$, we see it remains a partial order when restricted to some $\mathcal{P}(M)(A) \subseteq \mathcal{P}(A)$.

And this observation evidently generalizes to give us a trivial mini-theorem:

Theorem 1.2 *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. Again we could dignify it as a mini-theorem:

Theorem 1.3 *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* $y \preccurlyeq x$*.*

The proof is a trivial exercise. 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 $\leqslant$, 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 represented by $\subseteq$, the converse relation which orders its dual will naturally 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 called *weakly* ordered sets. By contrast,

Definition 1.2 *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* $y \not\prec x$*;*
(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 $\langle \mathbb{N}, \prec \rangle$ 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 take care with 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 two-part theorem, whose proof can again be left as an exercise.

Theorem 1.4 *Suppose* $\mathscr{P} = \langle P, \preccurlyeq \rangle$ *is a partially ordered set, and for* $x, y \in P$ *put* $x \prec y =_{\text{def}} x \preccurlyeq y \land x \neq y$*. Then* $\mathscr{P}^{-} = \langle P, \prec \rangle$ *is a strictly ordered set. Conversely, suppose* $\mathscr{Q} = \langle Q, \prec \rangle$ *is a strictly ordered set, and for* $x, y \in Q$ *put* $x \preccurlyeq y =_{\text{def}} x \prec y \lor x = y$*. Then* $\mathscr{Q}^{+} = \langle Q, \preccurlyeq \rangle$ *is a partially ordered set.*

When we have a partially ordered set we can therefore immediately define a counterpart strictly ordered set by ignoring identies, and vice versa. 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 trivial 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 to take the idea of posets as the fundamental one.

1.3 Elementary relations among posets

Next, we'll define three different, increasingly constrained, kinds of maps that can relate a pair of posets:

Definitions 1.3 *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* $f(p) \sqsubseteq f(p')$.
(iii) f *is an* order-isomorphism *iff* f *is a surjection (is onto) and is an order-embedding.*

The basic idea is that a monotone mapping f doesn't actually reverse orderings (in an obvious notation, if $p \prec p'$ we can't actually get $f(p') \sqsubset f(p)$), but it can squeeze up orderings (we can have $p \prec p'$ but $f(p) = f(p')$). An order-embedding however preserves more: if $p \prec p'$ then $f(p) \sqsubset f(p')$, and so the embedding puts a faithful copy of $\mathscr{P}$'s ordering inside $\mathscr{Q}$. While an order-isomorphism, as you'd expect from the label, makes the whole of $\mathscr{Q}$ an order-preserving copy of $\mathscr{P}$.

We need to check that last claim. For we are given that an order-isomorphism is a surjection. But to be a true isomorphism it also needs to be an injection (i.e. it needs to be a one-one mapping between the carrier sets). But it is, since it is an embedding, and we have:

Theorem 1.5 *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 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 that $p \preccurlyeq p'$ and $p' \preccurlyeq p$. And since $\preccurlyeq$ is anti-symmetric, it follows that $p = p'$.

So, in sum, if $f(p) = f(p')$, then $p = p'$; i.e. if $p \neq p'$ then $f(p) \neq f(p')$, so f is one-one. $\square$

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

Theorem 1.6 *If* $f\colon P \to Q$ *is an order-embedding between* $\mathscr{P} = \langle P, \preccurlyeq\rangle$ *and* $\mathscr{Q} = \langle Q, \preccurlyeq\rangle$*, then* f *is an order-isomorphism between* $\mathscr{P}$ *and the poset* $\langle f(P), \preccurlyeq\rangle$*.*

$f(P)$ is, of course, the image of P under f, i.e. $\{q \mid (\exists p \in P) f(p) = q\}$. And the abuse of notation comes in not explicitly marking that the order-relation in $\langle f(P), \preccurlyeq\rangle$ is the relation in $\langle Q, \preccurlyeq\rangle$ restricted to $f(P)$. But here and henceforth we'll allow ourselves that degree of slack. The proof is trivial.

Let's use these definitions 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).

Note again that a monotone map typically *doesn't* preserve all order information: that's why the other standard label for such a map, i.e. 'order-preserving', can mislead. For an extreme example, 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 trivially monotone, but it's a forgetful map that throws away *all* order information.

(2) Take the same pair of 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 $\mathbb{R}^+$.

(3) 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.

(4) But don't be misled by that last example – there isn't always an order-isomorphism between a poset and its 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 $\langle \mathbb{N}, \geqslant \rangle$.

It is important to note that monotone maps (and likewise order-embedding, order-isomorphic maps) can be compounded together to get more:

Theorem 1.7 *If* $f\colon P \to Q$ *is monotone map between* $\mathscr{P} = \langle P, \preccurlyeq \rangle$ *and* $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$*, and* $g\colon Q \to R$ *is a monotone map between* $\mathscr{Q}$ *and* $\mathscr{R} = \langle R, \leqslant \rangle$*, then 'g following f', i.e. the composite map* $g \circ f\colon P \to R$*, is a monotone map between* $\mathscr{P}$ *and* $\mathscr{R}$*.*

Similarly for order-embeddings and for order-isomorphisms.

Proof If $p \preccurlyeq p'$ then $f(p) \sqsubseteq f(p')$, since f is monotone. And $f(p) \sqsubseteq f(p')$ 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. The other cases are similar. □

1.4 Order types

We have seen that – as the label suggests – 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 1.4 *The posets* $\mathscr{P}$ *and* $\mathscr{Q}$ *are* order-similar*, or* have the same order-type *(in symbols,* $\mathscr{P} \cong \mathscr{Q}$*), iff there is an order-isomorphism from* $\mathscr{P}$ *to* $\mathscr{Q}$*.*

It is an easy result that similarity is indeed properly so-called:

Theorem 1.8 *Order-similarity is an equivalence relation between posets.*

Proof It is trivial that order-similarity is reflexive. And to see that it is symmetric, we just need to confirm that if $f\colon P \to Q$ between $\mathscr{P}$ and $\mathscr{Q}$ is an order-isomorphism, then (i) we can define an inverse one-one correspondence $f^{-1}\colon Q \to P$ such that $f^{-1}(q) = p$ iff $f(p) = q$, and (ii) this is an order-isomorphism between $\mathscr{Q}$ and $\mathscr{P}$.

And to prove the transitivity of order-similarity, we need to confirm Theorem 1.7's claim that (iii) if $f\colon P \to Q$ is an order-isomorphism between $\mathscr{P}$ and $\mathscr{Q}$, and $g\colon Q \to R$ is an order-isomorphism between $\mathscr{Q}$ and $\mathscr{R}$, then the composite function $g \circ f\colon P \to R$ is an order-isomorphism between $\mathscr{P}$ and $\mathscr{R}$.

Detailed proofs of (i) to (iii) can safely be left as exercises. □

Again, 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 $\langle 2\mathbb{N}, \leqslant \rangle$ are all order-similar (where $\mathbb{N} \setminus \{0\}$ is the set of naturals with 0 deleted, and $2\mathbb{N}$ 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 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 $q \in \mathbb{Q}^+$ one-to-one to the pair of naturals $\langle m, n\rangle$ where $q = m/n$, with the fraction is in lowest terms. Do a standard kind of zig-zag enumeration of the pairs $\langle m, n\rangle$ where m, n have no common divisors, and put $q \preccurlyeq q'$ in $\mathbb{Q}^+$ when 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) We have also already seen that $\langle \mathbb{R}, \leqslant\rangle \cong \langle \mathbb{R}, \geqslant\rangle$. Later we'll show that the poset $\langle \mathbb{R}^2, \preccurlyeq\rangle$, i.e. the ordered pairs of reals ordered lexicographically, isn't order-similar to either.

(4) These examples so far 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 aren't totally ordered but which are order-similar, consider the following. (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!

1.5 Maximal, maximum, supremum

Let's now add two more bunches of definitions, this time to do with – so to speak loosely – the tops and bottoms of posets or parts of posets.

Definitions 1.5 *Suppose $\langle P, \preccurlyeq\rangle$ is a poset. Then $x \in P$ is* maximal *iff for all $y \in P$, if $x \preccurlyeq y$ then $y = x$. And x is a* maximum *if for all $y \in P$, $y \preccurlyeq x$. More generally, if $X \subseteq P$, then x is a* maximum of X *if for all $y \in X$, $y \preccurlyeq x$.*

Dually, $x \in X$ is minimal *iff for all $y \in P$, if $y \leq x$ then $y = x$. And x is a* minimum *if for all $y \in P$, $x \preccurlyeq y$: x is a* minimum of X *if for all $y \in X$, $x \preccurlyeq y$.*

In other words, a poset has a maximal element if there's an element such that no distinct element comes after it in the ordering. Maximal elements need not be unique (for a trivial example, consider the poset $\langle \mathbb{N}, =\rangle$). But if a poset has a maximum, then that *is* unique (for if x and y are each maxima, we'd have both $y \preccurlyeq x$ and $x \preccurlyeq y$, and hence $x = y$).

Likewise, local maxima in subsets are unique when they exist. Dually, of course, for minimals and minima.

Some examples:

(1) $\langle \mathbb{N}, \leqslant \rangle$ has a minimum and no maximum, though every *finite* subset $X \subseteq \mathbb{N}$ has a maximum; $\langle \mathbb{N}, \geqslant \rangle$ has a maximum and no minimum (careful!); $\langle \mathbb{R}, \leqslant \rangle$ has neither.

(2) $\langle \mathbb{N} \setminus \{0\}, \preccurlyeq \rangle$, where $\preccurlyeq$ represents ordering by divisibility, has a minimum – for the number 1 divides all numbers. $\langle \mathbb{N} \setminus \{0, 1\}, \preccurlyeq \rangle$ on the other hand has no minimum, though every prime number is minimal.

(3) The special inclusion powerset $\langle \mathcal{P}(A), \supseteq \rangle$ has a minimum and a maximum. But that's *not* typical of inclusion posets.

(4) It is immediate that a poset with a maximum element has a unique maximal element. What about the reverse? It is momentarily tempting to think that a unique maximal element would have to be a maximum. But not so. Take, for example, the set of natural numbers $\mathbb{N}$ and add a rogue element α. And define the following order on this set: for $m, n \in \mathbb{N}$, $m \preccurlyeq n$ iff $m \leqslant n$; and otherwise $0 \preccurlyeq \alpha$, and $\alpha \preccurlyeq \alpha$. Then α is a unique maximal element in the poset $\langle \mathbb{N} \cup \{\alpha\}, \preccurlyeq \rangle$, but isn't a maximum.

Definitions 1.6 *Suppose again that $\langle P, \preccurlyeq \rangle$ is a poset, and let $X \subseteq P$. Then $x \in P$ is an* upper bound *for X iff for all $y \in X$, $y \preccurlyeq x$. And x is a* supremum *(or* least upper bound*) of X, if x is an upper bound for X and, for all y, if y is an upper bound of X, $x \preccurlyeq y$.*

Dually, $x \in P$ is a lower bound *for X iff for all $y \in X$, $x \preccurlyeq y$. And x is a* infimum *(or* greatest lower bound*) of X, if x is a lower bound for X and, if for all y, if y is a lower bound of X, $y \preccurlyeq x$.*

Evidently, if X has a supremum, it is unique (for if x and y are both *least* upper bounds, then $x \preccurlyeq y$ and $y \preccurlyeq x$ hence $x = y$). Further, if X has a maximum, then that same element is its supremum (check the definitions!); but – as we'll see – a set can have a supremum even if it lacks a maximum. Dually, of course, for infima and minima. A few examples for our latest definitions:

(5) In $\langle \mathbb{N}, \leqslant \rangle$, every subset $X \subseteq \mathbb{N}$ has an infimum; but infinite subsets lack an upper bound (are 'open at the top'); but any which has an upper bound has a least upper bound – and the least upper bound is its maximum.

(6) In $\langle\mathbb{Q}+, \leqslant\rangle$, again every subset $X \subseteq \mathbb{N}$ has an infimum and some are 'open at the top'. However, this time there are subsets of the positive rationals which are bounded above but which have no least upper bound (and hence have no maxima). For example, take the set of positive rationals q which are such that $q^2 \leq 2$; this is bounded above, but since $\sqrt{2}$ is not rational, lacks a supremum in the rationals. However, if you equip $\mathbb{Q}+$ with the 'zig-zag' ordering to get the poset $\langle\mathbb{Q}^+, \preccurlyeq\rangle$, then every subset bounded above does have a supremum.

(7) It is a key proposition of classical analysis that, in $\langle\mathbb{R}, \leqslant\rangle$, every subset bounded above has a supremum (and dually of course). But a subset bounded above need not have a maximum. Take, for example, the set of reals $\{x \mid x < \sqrt{2}\}$ which contains no maximum element but whose supremum – *outside* the set of course! – is $\sqrt{2}$.

(8) In the special inclusion poset $\langle\mathcal{P}(A), \subseteq\rangle$, every subset of $\mathcal{P}(A)$ has a maximum and hence supremum (the union) and a minimum and infimum (the intersection). But this won't in general be the case for inclusion posets.

Let's just note one portmanteau theorem about these notions.

Theorem 1.9 *If* $f\colon P \to Q$ *is a monotone mapping from* $\mathscr{P} = \langle P, \preccurlyeq\rangle$ *to* $\mathscr{Q} = \langle Q, \sqsubseteq\rangle$*, it sends elements which are maximal in* P *to elements maximal in* Q*, sends maxima to maxima, local maxima to local maxima, upper bounds to upper bounds, and suprema to suprema. Likewise for their minima, etc.*

The proof is immediate by inspection of cases. And since the result applies to all monotone mappings it applies to embeddings and isomorphisms in particular.

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 1.10 *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 actually 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.

Indeed, if we really, *really*, want to, we can concentrate on inclusion posets whose carrier set itself contains just *pure sets* as its elements. Start again from $\mathscr{P} = \langle P, \preccurlyeq\rangle$. Construct an isomorphic poset $\mathscr{M} = \langle S, \sqsubseteq\rangle$ by the brute force method of (i) mapping each distinct element $p \in P$ one-to-one to a distinct pure set $s(p)$ – there should be enough sets to go around! – and then (ii) putting S to be the set of such correlates and defining $s(p) \sqsubseteq s(p')$ to hold iff $p \preccurlyeq p'$. Trivially, $\mathscr{P} \cong \mathscr{M}$, and by the same construction as before $\mathscr{M} \cong \mathscr{M}^\pi$, where $\mathscr{M}^\pi$ is now an inclusion poset whose carrier set contains just pure sets. Hence $\mathscr{P} \cong \mathscr{M}^\pi$. (It is, of course, the possibility of this sort of trickery done on the grand scale that leads people to say that we can reflect all – or nearly all – mathematics inside pure set theory.)

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, and we are going to move on rather rapidly to talk about ordered sets which have a bit of additional structure.

Still, before we do so, it *is* perhaps worth pausing to explore another common relation that can obtain between posets, one which is weaker than order-isomorphism but which is surprisingly rich in implications. True, this will be something of an early digression from our main path. But the ideas are cute, and 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 *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_*.[1]

---

[1] 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.

Talk of adjoints is carried over from category theory. Category theorists in turn seem to have borrowed the term from the old theory of Hermitian operators, where

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 don't want to get bogged down in unnecessary mathematics; so for the moment let's just give some very simple cases.

(1) Suppose f is an order-isomophism 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) Our third mini-example is from elementary logic (though we can give an parallel example in any Boolean algebra).

Recall example 12 from Section 1.1. So L is a propositional language; $|\alpha|$ is the class of L-wffs logically equivalent to α; E is the set of all such equivalence classes; and $|\alpha| \;\;\Rightarrow\!|\;\; \beta| \in E$ iff $\alpha \vDash \beta$. Then $\langle E, \Rightarrow\rangle$ is a poset.

Now consider the following two functions between E and itself. Fix γ to be some L-wff. Then put f_* to be the function that maps the equivalence class $|\alpha|$ to the class $|(\gamma \wedge \alpha)|$, and f^* is the function that maps the equivalence class $|\alpha|$ to the class $|(\gamma \supset \alpha)|$. 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.')

Our first family of cases shows that Galois connections are at least as plentiful as order-isomorphisms. The second case shows that posets that

---

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).

aren't order-isomorphic can still be connected. And our third toy case shows that even when the Galois connected posets *are* isomorphic (in this case because they are identical!), the functions f_* and f^* that go to make up the connection needn't be isomorphisms.

2.2 An alternative definition

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

Theorem 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 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 $p \preccurlyeq p'$. Then since we've just 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. $\Box$

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]

---

[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 $p \preccurlyeq f^*(q)$.

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 2.2 *If* $\langle f_*, f^1 \rangle$ *and* $\langle f_*, f^2 \rangle$ *are both 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 both Galois connections between the same posets, then* $f_1 = f_2$*.*

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 2.1 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 2.1 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 exist a unique corresponding f^* such that $\langle f_*, f^* \rangle$ form a Galois connection (we'll find a counterexample 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 give you the materials to see 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 *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* $\{q \in Q \mid p \preccurlyeq f^*(q)\}$

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), but (b) 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 To show (i), 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 2.1, $p \preccurlyeq f^*(f_*(p))$, so $p \preccurlyeq f^*(q)$. Hence $f^*(q)$ is certainly an upper bound to $\{p \in P \mid f_*(p) \sqsubseteq q\}$. To show it is a maximum, we need to check that $f^*(q) \in P$ is actually in that set. But it is because, by Theorem 2.1 again, $f_*(f^*(q)) \sqsubseteq q$.

The proof of (ii) 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 an 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 2.4 *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 that that same 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 $f(A)$ for the image of the set A under f, i.e. $f(A) = \{f(a) \mid a \in A\}$. And a is a fixed point of a function f if $f(a) = a$. Then we have the following:

Theorem 2.5 *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 (i) 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 (i).

(ii) 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 $q \in f_*(P)$.

(iii) 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 (iii). ☐

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

Definition 2.2 *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 2.6 *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 2.5, $\mathscr{P}^f = \langle f^*(Q), \preccurlyeq \rangle$ 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 2.3 *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 2.7 *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 (i) $p \preccurlyeq f^* \circ f_*\,(p)$ by Theorem 2.1.

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

(iii) Theorem 2.5 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 been talking about Galois connections in general. Let's now note one characteristic way by which some connections can be generated.

Theorem 2.8 *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.

To spell things out, (i) $f_R(\alpha) \supseteq \beta$ unpacks as (ii), for every $b \in \beta$, $\forall a(a \in \alpha \to aRb)$. Rearranging, (ii) equivalent to (iii): for every $a \in \alpha$, $\forall b(b \in \beta \to aRb)$. Which can be abbreviated as (iv), $\alpha \subseteq f^R(\beta)$. □

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 γ and δ – is *relation-generated*. Galois's original example was of this kind. The example we will be looking at in the rest of this chapter is relation-generated too.

2.6 Syntax and semantics

We turn – at last! – to consider a simple but central example which makes Galois connections of interest to logicians.

In his classic 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 to category theorists, the idea still doesn't seem to be that widely known. So let's explain and explore. We'll start in this section with a bunch of reminders (or you can just skip to the definition at the end).

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. So the claim is that we can make a Galois connection between sets of sentences and classes of models.

To keep things under control, let's think initially about sets of sentences drawn from a classical first-order language. We can take it, then, that L has a standard *logical* vocabulary, i.e. an adequate set of propositional connectives including classical negation, a pair of quantifier formers $\forall$ and $\exists$ and an unlimited supply of variables, an identity symbol, and parentheses. The particular choice of logical symbolism isn't significant. And then L has its distinctive *non-logical* vocabulary. That is to say, L has a certain fixed number of constants; a certain fixed number of predicates, each with a given arity; and a certain fixed number of L-function-symbols, each with a given arity. (Any of the sets of constants, predicates and functions might be empty.)

For example, if L_A is a suitable language for first-order Peano Arithmetic, it has the standard logical vocabulary, and it comes equipped with a single constant '$\mathsf{0}$', perhaps a single two-place predicate '$\leqslant$', and three function symbols, a one-place function-symbol 'S', and the two-place function symbols '$+$' and '$\times$'.

The basic rules for forming closed sentences from our language's logical and non-logical vocabulary are then familiar: we needn't fuss about the fine print.

Given a language L, an L-structure s is one apt for interpreting the language. So the structure has a non-empty domain. And then corresponding to each L-constant, s provides a designated element of the domain. Corresponding to each L predicate, s provides a relation defined over the domain and which has the same arity as the predicate. And corresponding to each L function symbol, s provides a (total) function with the matching arity.[3]

For example, one suitable structure (the 'standard model') for L_A takes the domain to be the natural numbers, assigns zero to the constant '$\mathsf{0}$', the less-than-or-equals relation to '$\leqslant$', the monadic successor function to 'S', and the dyadic addition and multiplication functions to '$+$' and '$\times$'.

We'll take the general idea of a structure apt to interpret a first-order language L to be very familiar. And we can then define what it is for an L-sentence φ to be true with respect to an L-structure s in the entirely

---

[3] Of course, the particular choice of non-logical symbolism isn't logically significant. Abstract away from the particular symbolism used by L and we are left with the *signature* of the language. And it is this abstract signature (not, the particular choice of symbolism) that really determines the shape of an appropriate structure which we can use to interpret the language. But for brevity, we'll continue to talk of 'L-structures' rather than talk, as perhaps we should, of structures matching the signature of L.

familiar way. When s indeed makes φ true, we write '$s \vDash \varphi$' (that's a standard overloading of symbolism).

Just one more reminder in this preamble. We want, finally, the notion of a model, and it's worth highlighting the definition:

Definition 2.4 *The L-structure s is a* model *of a set of L-sentences α iff for every $\varphi \in \alpha$, $s \vDash \varphi$.*

2.7 Making the connection

Now we make Lawvere's connection. On the same assumption that L is a (first-order) language,

Definitions 2.5 *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 2.8, it is immediate that

Theorem 2.9 *$\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{A}$ and $\mathscr{S}$ (the connection generated by the converse of the relation $\vDash$, which holds between elements of A and S).*

2.8 Drawing the consequences

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

Theorem 2.10 *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. But what do these results really *mean*?

Result (i) just tells 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.)

Results (ii) to (iv) are more interesting. For 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 perhaps 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 2.6 *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, $\theta(\alpha)$, is the smallest theory containing α.*[4]

---

[4] 'Theories' are, of course, equally often defined as sets of sentences closed under some syntactic deducibility relation (rather than as sets closed under semantic

We consequently have

Theorem 2.11 *$f^*(f_*(\alpha))$ 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 2.10 (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 2.10 (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 2.10 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 2.10? 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 $f_*(f^*(\mu))$ is the set of *all* structures which are models for that theory. Which links up with another familiar idea:

Definition 2.7 *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

---

consequence). So let's emphasize: it is the semantic relation that is in play in our definition here.

Theorem 2.12 *$f_*(f^*(\mu))$ is an axiomatizable class of structures, and it is the smallest axiomatizable class containing μ.*

Proof Dual to the proof of Theorem 2.11 $\square$

And looked at through the lens of *this* theorem, the last three parts of Theorem 2.10 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.

2.9 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 reaction, but one that rather misses the point. So two comments.

First, the appearance of mere triviality comes when we look at Theorem 2.10 in the light of Theorems 2.11 and 2.12. 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 framework. Look at it this way. Start from the *true-of* relation which can contain between an L-sentence and an L-structure. That immediately generates in a natural way a Galois connection between two posets whose elements are now sets of sentences and sets of strctures. 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 kinds of mathematics aren't really for you!

2.10 More consequences

It will perhaps help us to see better what's going on here when we consider the isomorphism Theorem 2.6 which implies that

Theorem 2.13 *The subposets* $\mathscr{A}^f = \langle f^*(f_*(\mathcal{P}(A))), \subseteq\rangle$ *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, but *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 (consistent) theories, on to more and more partial (consistent) 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.

2.11 Posets as sets

And I think that's about as much juice as we can usefully squeeze out of the simple observation that $\langle f_*, f^* \rangle$ as defined is a Galois connection between sets of L-sentences, and sets of L-structures. We could generalize a bit in various directions, by fiddling with the various bits of our assumption that L is a first-order classical language of fixed signature: but we won't learn anything very exciting by so doing. So we won't.

But we ought to finish by making one general comment, if only to quiet the worry that might have occurred to you when confronted with Definitions 2.5.

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, 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.[5]

However, we *have* been helping ourselves throughout to the idiom of set-talk all the way through, 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 rather perverse to go out of our way to avoid it. The extra

---

[5] Of course, as we saw in Section 1.6, if we *do* countenance sets we'll be able to relate other pluralities of objects and their partial orderings with isomorphic set-theoretic constructs: the present point is only that thinking about objects-with-orderings doesn't yet *have* to involve thinking of sets.

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 an L-structure, if we 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's going to as big as the universe of all sets, period. So the carrier set for $\mathscr{S}$, on standard views, too big to be a kosher set. It's a '(proper) class', which is no doubt why Lawvere talked 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 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 using a plural idiom instead (so, for example, we think of f_* not as mapping a set of sentences to the set or class of their models, but as mapping some sentences to their models). But the point remains that, even when using set talk as we have done, we don't ever need to lean hard on the presumption that all the collections of L-structures are themselves objects which can be collected into a new object, the oversized Set $\mathcal{P}(S)$. Talking of $\mathscr{S}$ as a poset as we did, at least in the present context, is just a convenience, which we can insouciantly allow ourselves.