Partial Orders, Lattices, Equivalence Relations

1. Partial Orders

There are two main kinds of relations that play a very important role in mathematics and computer science:

1. Partial orders
2. Equivalence relations.

In this section and the next few ones, we define partial orders and investigate some of their properties.

Given a set, X, we can order the subsets of X by the subset relation: $A \subseteq B$, where A, B are any subsets of X.

For example, if $X = \{a, b, c\}$, we have $\{a\} \subseteq \{a, b\}$. However, note that neither $\{a\}$ is a subset of $\{b, c\}$ nor $\{b, c\}$ is a subset of $\{a\}$.

We say that $\{a\}$ and $\{b, c\}$ are *incomparable*.

Not all relations are partial orders, so which properties characterize partial orders?

Definition 1: A binary relation, $\leq$, on a set, X, is a *partial order* (or *partial ordering*) if it is *reflexive*, *transitive* and *antisymmetric*, that is:

(1) (*Reflexivity*): $a \leq a$, for all $a \in X$;

(2) (*Transitivity*): If $a \leq b$ and $b \leq c$, then $a \leq c$, for all $a, b, c \in X$.

(3) (*Antisymmetry*): If $a \leq b$ and $b \leq a$, then $a = b$, for all $a, b \in X$.

A partial order is a *total order (ordering)* (or *linear order (ordering)*) if for all $a, b \in X$, either $a \leq b$ or $b \leq a$.

When neither $a \leq b$ nor $b \leq a$, we say that *a and b are incomparable*.

A subset, $C \subseteq X$, is a *chain* if $\leq$ induces a total order on C (so, for all $a, b \in C$, either $a \leq b$ or $b \leq a$).

The *strict order (ordering), $<$, associated with $\leq$* is the relation defined by: $a < b$ iff $a \leq b$ and $a \neq b$.

If $\leq$ is a partial order on X, we say that the pair $(X, \leq)$ is a *partially ordered set* or for short, a *poset*.

Remark: Observe that if $<$ is the strict order associated with a partial order, $\leq$, then $<$ is transitive and *anti-reflexive*, which means that

(4) $a <> a$, i.e., not comparable, for all $a \in X$.

Conversely, let $<$ be a relation on X and assume that $<$ is transitive and anti-reflexive.

Then, we can define the relation $\leq$ so that $a \leq b$ if $a = b$ or $a < b$.

It is easy to check that $\leq$ is a partial order and that the strict order associated with $\leq$ is our original relation, $<$.

Given a poset, $(X, \leq)$, by abuse of notation, we often refer to $(X, \leq)$ as the *poset X*, the partial order $\leq$ being implicit.

If confusion may arise, for example when we are dealing with several posets, we denote the partial order on X by $\leq_X$.

Here are a few examples of partial orders.

1. The subset ordering. We leave it to the reader to check that the subset relation, $\subseteq$, on a set, X, is indeed a partial order.

   For example, if $A \subseteq B$ and $B \subseteq A$, where $A, B \subseteq X$, then $A = B$, since these assumptions are exactly those needed by the extensionality axiom.

2. The natural order on $\mathbb{N}$.

3. Orderings on strings.

   Let $\Sigma = \{a_1, \ldots, a_n\}$ be an alphabet. The prefix substring relation is a partial order.

Given a poset, $(X, \leq)$ if X is finite, then there is a convenient way to describe the partial order $\leq$ on X using a graph.

Consider an arbitrary poset, $X \leq$ (not necessarily fi- nite). Given any element, $a \in X$, the following situations are of interest:

1. For no $b \in X$ do we have $b < a$. We say that a is a *minimal element* (of X).
2. There is some $b \in X$ so that $b < a$ and there is no $c \in X$ so that $b < c < a$. We say that b is an *immediate predecessor of a*.
3. For no $b \in X$ do we have $a < b$. We say that a is a *maximal element* (of X).
4. There is some $b \in X$ so that $a < b$ and there is no $c \in X$ so that $a < c < b$. We say that b is an *immediate successor of a*.

Note that an element may have more than one immediate predecessor (or more than one immediate successor).

If X is a finite set, then it is easy to see that every element that is not minimal has an immediate predecessor and any element that is not maximal has an immediate successor (why?).

But if X is infinite, for example, $X = \mathbb{Q}$, this may not be the case. Indeed, given any two distinct rational numbers, $a, b \in \mathbb{Q}$, we have

$$a < \frac{a+b}{2} < b.$$

Let us now use our notion of immediate predecessor to draw a diagram representing a finite poset, $(X, \leq)$

The trick is to draw a picture consisting of nodes and oriented edges, where the nodes are all the elements of X and where we draw an oriented edge from a to b iff a is an immediate predecessor of b.

Such a diagram is called a *Hasse diagram* for $(X, \leq)$

Observe that if $a < c < b$, then the diagram does not have an edge corresponding to the relation $a < b$.

A Hasse diagram is an economical representation of a fi- nite poset and it contains the same amount of information as the partial order, $\leq$.

Here is the diagram associated with the partial order on the power set of the two element set, $\{a, b\}$:

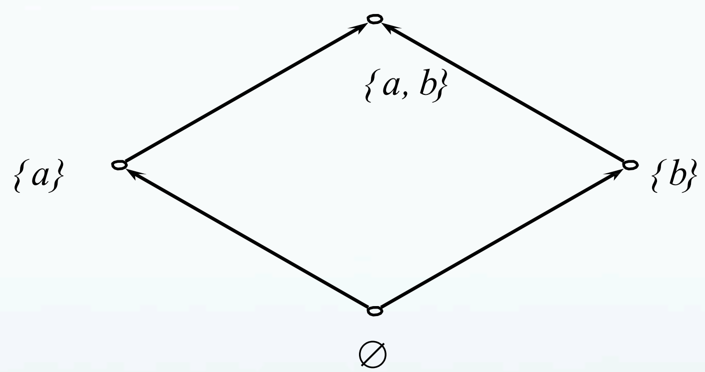

Figure 1: The partial order of the power set $2^{\{a,b\}}$

Here is the diagram associated with the partial order on the power set of the three element set, $\{a, b, c\}$:

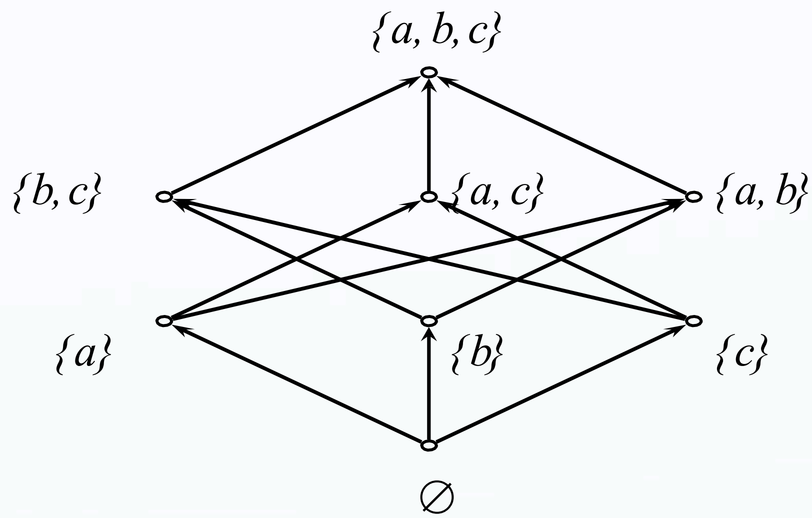

Figure 2: The partial order of the power set $2^{\{a,b,c\}}$

Note that $\emptyset$ is a minimal element of the above poset (in fact, the smallest element) and $\{a, b, c\}$ is a maximal el- ement (in fact, the greatest element).

In the above example, there is a unique minimal (resp. maximal) element.

A less trivial example with multiple minimal and maximal elements is obtained by deleting $\emptyset$ and $\{a, b, c\}$:

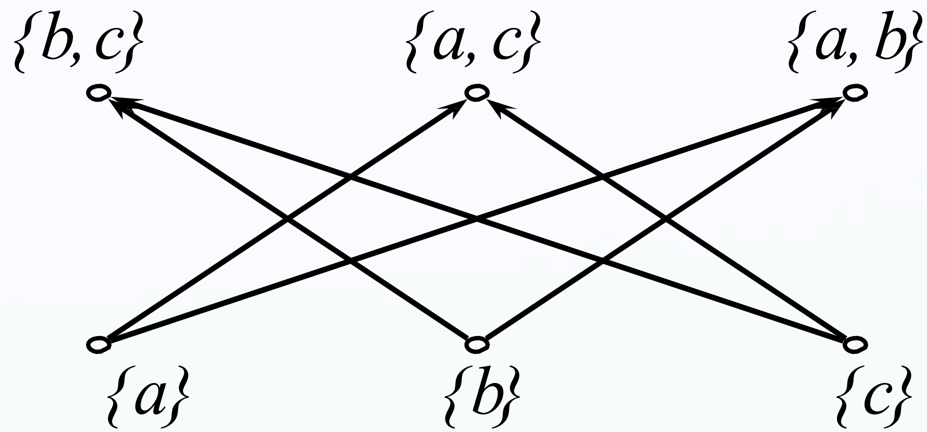

Figure 3: Minimal and maximal elements in a poset

Given a poset, $X, \leq$ observe that if there is some ele- ment $m \in X$ so that $m \leq x$ for all $x \in X$, then m is unique.

Such an element, m, is called the *smallest* or the *least element* of X.

Similarly, an element, $b \in X$, so that $x \leq b$ for all $x \in X$ is unique and is called the *greatest element* of X.

Definition 2: Let $(X, \leq)$ be a poset and let $A \subseteq X$ be any subset of X. An element, $b \in X$, is a *lower bound of A* iff $b \leq a$ for all $a \in A$.

An element, $m \in X$, is an *upper bound of A* iff $a \leq m$ for all $a \in A$.

An element, $b \in X$, is the *least element of A* iff $b \in A$ and $b \leq a$ for all $a \in A$.

An element, $m \in X$, is the *greatest element of A* iff $m \in A$ and $a \leq m$ for all $a \in A$.

An element, $b \in A$, is *minimal in A* iff $a < b$ for no $a \in A$, or equivalently, if for all $a \in A$, $a \leq b$ implies that $a = b$.

An element, $m \in A$, is *maximal in A* iff $m < a$ for no $a \in A$, or equivalently, if for all $a \in A$, $m \leq a$ implies that $a = m$.

An element, $b \in X$, is the *greatest lower bound of A* iff the set of lower bounds of A is nonempty and if b is the greatest element of this set.

An element, $m \in X$, is the *least upper bound of A* iff the set of upper bounds of A is nonempty and if m is the least element of this set.

Remarks:

1. If b is a lower bound of A (or m is an upper bound of A), then b (or m) may not belong to A.
2. The least element of A is a lower bound of A that also belongs to A and the greatest element of A is an upper bound of A that also belongs to A.
   When $A = X$, the least element is often denoted as ⊥, sometimes 0, and the greatest element is often denoted as T, sometimes 1.
3. Minimal or maximal elements of A belong to A but they are not necessarily unique.
4. The greatest lower bound (or the least upper bound) of A may not belong to A. We use the notation $\wedge A$ for the greatest lower bound of and the notation $\vee A$ for the least upper bound of .

   When $A = \{a, b\}$, we write $a \wedge b$ for $\wedge\{a, b\}$ and $a \vee b$ for $\vee\{a, b\}$.

   The element $a \wedge b$ is called the *meet of a and b* and $a \vee b$ is the *join of a and b*. (Some computer scientists use $a \quad b$ for $a \wedge b$ and $a \quad b$ for $a \vee b$.)
5. Observe that if it exists, $\wedge X = \perp$ and if it exists, $\vee X =$ T.

Figure 4: Max Zorn, 1906-1993

Zorn's lemma turns out to be equivalent to the axiom of choice.

Theorem 1 (*Zorn's Lemma*) *Given a poset,* $(X, \leq)$ *if every nonempty chain in* X *has an upper bound, then* X *has some maximal element.*

When we deal with posets, it is useful to use functions that are order-preserving as defined next.

Definition 3: Given two posets $(X, \leq_X)$ and $(Y, \leq_Y)$, a function, $f : X \to Y$, is *monotonic* (or *order- preserving*) if for all $a, b \in X$, $a \leq_X$ implies that $f(a) \leq_Y f(b)$.

Lattices and Tarski's Fixed Point Theorem

We now take a closer look at posets having the property that every two elements have a meet and a join (a greatest lower bound and a least upper bound).

Such posets occur a lot more than we think. A typical example is the power set under inclusion, where meet is intersection and join is union.

Definition 4: A *lattice* is a poset in which any two elements have a meet and a join. A *complete lattice* is a poset in which any subset has a greatest lower bound and a least upper bound.

According to part (5) of the remark just before Zorn's Lemma, observe that a complete lattice must have a least element, $\perp$, and a greatest element, T.

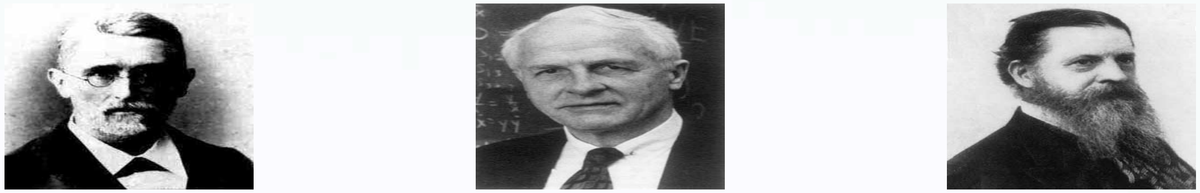

Figure 5: J.W. Richard Dedekind, 1831-1916 (left), Garrett Birkhoff, 1911-1996 (middle) and Charles S. Peirce, 1839-1914 (right)

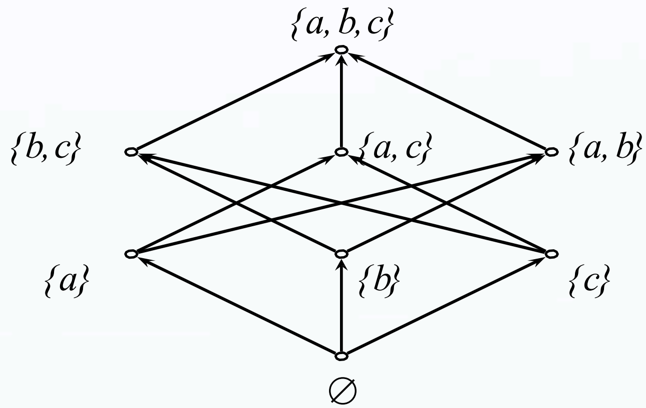

Figure 6: The lattice $2^{\{a,b,c\}}$

Remark: The notion of complete lattice is due to G. Birkhoff(1933). The notion of a lattice is due to Dedekind (1897) but his definition used properties (L1)-(L4) listed in Proposition 5.2.2. The use of meet and join in posets was first studied by C. S. Peirce (1880).

Figure 5.6 shows the lattice structure of the power set of $\{a, b, c\}$. It is a complete lattice.

It is easy to show that any finite lattice is a complete lattice and that a finite poset is a lattice iff it has a least element and a greatest element.

The poset $\mathbb{N}_+$ under the divisibility ordering is a lattice! Indeed, it turns out that the meet operation corresponds to *greatest common divisor* and the join operation corresponds to *least common multiple*.

However, it is not a complete lattice.

The power set of any set, X, is a complete lattice under the subset ordering.

Proposition 1: *If X is a lattice, then the following identities hold for all $a, b, c \in X$:*

$L1$	$a \vee b = b \vee a,$	$a \wedge b = b \wedge a$
$L2$	$(a \vee b) \vee c = a \vee (b \vee c),$	$(a \wedge b) \wedge c = a \wedge (b \wedge c)$
$L3$	$a \vee a = a,$	$a \wedge a = a$
$L4$	$(a \vee b) \wedge a = a,$	$(a \wedge b) \vee a = a.$

Properties (L1) correspond to commutativity, prop- erties (L2) to associativity, properties (L3) to idem- potence and properties (L4) to absorption. Furthermore, for all $a, b \in X$, we have

$$a \leq b \qquad \text{iff } a \vee b = b \qquad \text{iff } a \wedge b = a,$$

called consistency.

Properties (L1)-(L4) are algebraic properties that were found by Dedekind (1897).

A pretty symmetry reveals itself in these identities: they all come in pairs, one involving $\wedge$, the other involving $\vee$.

A useful consequence of this symmetry is *duality*, namely, that each equation derivable from (L1)-(L4) has a dual statement obtained by exchanging the symbols $\wedge$ and $\vee$.

What is even more interesting is that it is possible to use these properties to define lattices.

Indeed, if X is a set together with two operations, $\wedge$ and $\vee$, satisfying (L1)-(L4), we can define the relation $a \leq b$ by $a \vee b = b$ and then show that $\leq$ is a partial order such that $\wedge$ and $\vee$ are the corresponding meet and join.

Proposition 2: *Let X be a set together with two operations $\wedge$ and $\vee$ satisfying the axioms (L1)-(L4) of proposition 5.2.2. If we define the relation $\leq$ by $a \leq b$ iff $a \vee b = b$ (equivalently, $a \wedge b = a$), then $\leq$ is a partial order and $(X, \leq)$ is a lattice whose meet and join agree with the original operations $\wedge$ and $\vee$.*

Figure 7: Alferd Tarksi, 1902-1983

The following proposition shows that the existence of arbitrary least upper (or greatest lower) bounds is already enough ensure that a poset is a complete lattice.

Proposition 3: *Let $(X, \leq)$ be a poset. If X has a greatest element, $\top$, and if every nonempty subset, A, of X has a greatest lower bound, $\wedge A$, then X is a complete lattice. Dually, if X has a least element, $\bot$, and if every nonempty subset, A, of X has a least upper bound, $\vee A$, then X is a complete lattice.*

We now state a remarkable result due to A. Tarski (discovered in 1942, published in 1955).

Definition 5: Let $X, \leq$ be a poset and let $f : X \to X$ be a function. An element, $x \in X$, is a *fixed point of f* (sometimes spelled *fixpoint*) iff

$$f(x) = x.$$

An element, $x \in X$, is a *least (resp. greatest) fixed point of f* if it is a fixed point of f and if $x \leq y$ (resp. $y \leq x$) for every fixed point y of f.

Fixed points play an important role in certain areas of mathematics (for example, topology, differential equations) and also in economics because they tend to capture the notion of stability or equilibrium.

Theorem 2 (*Tarski's Fixed Point Theorem*) *Let* $(X, \leq)$ *be a complete lattice and let* $f : X \to X$ *be any monotonic function. Then, the set,* F*, of fixed points of* f *is a complete lattice. In particular,* f *has a least fixed point,*

$$x_{\min} = \bigwedge\{x \in X \mid f(x) \leq x\}$$

and a greatest fixed point

$$x_{\max} = \bigvee\{x \in X \mid x \leq f(x)\}.$$

Note that the least upper bounds and the greatest lower bounds in F do not necessarily agree with those in X. In technical terms, F is generally not a sub-lattice of X.

Tarski's Fixed Point Theorem can be used to prove the Schröder-Bernstein Theorem.

Theorem 3: *Given any two sets, A and B, if there is an injection from A to B and an injection from B to A, then there is a bijection between A and B.*

The proof is probably the shortest known proof of the Schröder-Bernstein Theorem because it uses Tarski's fixed point theorem, a powerful result.

If one looks carefully at the proof, one realizes that there are two crucial ingredients:

1. The set C is closed under $g \circ f$, that is, $g \circ f(C) \subseteq C$.
2. $A - C \subseteq g(B)$.

Using these observations, it is possible to give a proof that circumvents the use of Tarski's theorem. Such a proof is given in Enderton [1].

[1] Herbert B. Enderton. Elements of Set Theory. Academic Press, first edition, 1977.

Equivalence Relations and Partitions

Equivalence relations basically generalize the identity relation.

Technically, the definition of an equivalence relation is obtained from the definition of a partial order (Definition 1) by changing the third condition, antisymmetry, to *symmetry*.

Definition 6: A binary relation, R, on a set, X, is an *equivalence relation* if it is *reflexive*, *transitive* and *symmetric*, that is:

(1) (*Reflexivity*): aRa, for all $a \in X$;

(2) (*Transitivity*): If aRb and bRc, then aRc, for all $a, b, c \in X$.

(3) (*symmetry*): If aRb, then bRa, for all $a, b \in X$.

Here are some examples of equivalence relations.

1. The identity relation, id_X, on a set X is an equivalence relation.
2. The relation $X \times X$ is an equivalence relation.
3. Let S be the set of students in EE6226. Define two students to be equivalent iff they were born in the same year. We can check that this relation is an equivalence relation.
4. Given any natural number, $p \geq 1$, recall that we can define a relation on Z as follows:

$$n \equiv m \pmod p$$

   iff $p \mid n - m$, *i.e.*, $n = m + pk$, for some $k \in Z$. It is an easy exercise to check that this is indeed an equivalence relation called *congruence modulo p*.
5. Equivalence of propositions is the relation defined so that $P \equiv Q$ iff $P \Rightarrow Q$ and $Q \Rightarrow P$ are both provable (say, classically). The logical equivalence is an equivalence relation.
6. Suppose $f : X \to Y$ is a function. Then, we define the relation $\equiv_f$ on X by

$$x \equiv_f y \qquad \text{iff} \quad f(x) = f(y).$$

   It is immediately verified that $\equiv_f$ is an equivalence relation. We are going to show that every equivalence relation arises in this way, in terms of (surjective) functions.

The crucial property of equivalence relations is that they *partition* their domain, X, into pairwise disjoint nonempty blocks.

Definition 7: Given an equivalence relation, R, on a set, X, for any $x \in X$, the set

$$[x]_R = \{y \in X \mid xRy\}$$

is the *equivalence class of x*. Each equivalence class, $[x]_R$, is also denoted x_R and the subscript R is often omitted when no confusion arises. The set of equivalence classes of R is denoted by X/R. The set X/R is called the *quotient of X by R* or *quotient of X modulo R*. The function, $\pi : X \to X/R$, given by

$$\pi(x) = [x]_R, \qquad x \in X,$$

is called the *canonical projection* (or *projection*) of X onto X/R.

Since every equivalence relation is reflexive, *i.e.*, xRx for every $x \in X$, observe that $x \in [x]_R$ for any $x \in R$, that is, every equivalence class is *nonempty*.

It is also clear that the projection, $\pi : X \to X/R$, is surjective.

The main properties of equivalence classes are given by

Proposition 4: *Let R be an equivalence relation on a set, X. For any two elements $x, y \in X$, we have*

$$xRy \quad iff \quad [x] = [y].$$

Moreover, the equivalences classes of R satisfy the following properties:

(1) $[x] \neq \emptyset$ *for all* $x \in X$;

(2) If $[x] = [y]$ *then* $[x] \cap [y] = \emptyset$,

(3) $X = \bigcup_{x \in X} [x]$.

A useful way of interpreting Proposition 4 is to say that the equivalence classes of an equivalence relation form a partition, as defined next.

Definition 8: Given a set, X, a *partition of X* is any family, $\Pi = \{X_i\}_{i \in I}$, of subsets of X such that

(1) $X_i \neq \emptyset$ for all $i \in I$ (each X_i is nonempty);

(2) If $i \neq j$ then $X_i \cap X_j = \emptyset$ (the X_i are pairwise disjoint);

(3) $X = \bigcup_{i \in I} X_i$ (the family is exhaustive). Each set X_i is called a *block* of the partition.

In the example where equivalence is determined by the same year of birth, each equivalence class consists of those students having the same year of birth.

Let us now go back to the example of congruence modulo p (with $p > 0$). Can you tell what are the blocks of the corresponding partition? Recall that

$$m \equiv n \pmod p \text{ iff } m - n = pk \text{ for some } k \in \mathbb{Z}.$$

Given any set, X, let $\mathrm{Equiv}(X)$ denote the set of equivalence relations on X and let $\mathrm{Part}(X)$ denote the set of partitions on X.

Then, Proposition 4 defines the function, Π: $\mathrm{Equiv}(X) \to \mathrm{Part}(X)$, given by,

$$\Pi(R) = X/R = \{[x]_R \mid x \in X\},$$

where R is any equivalence relation on X. We also write Π_R instead of $\Pi(R)$.

There is also a function, $R : \mathrm{Part}(X) \to \mathrm{Equiv}(X)$, that assigns an equivalence relation to a partition, shown by the next proposition.

Proposition 5: *For any partition,* $\Pi = \{X_i\}_{i \in I}$, *on a set,* X, *the relation,* $R(\Pi)$, *defined by*

$$xR(\Pi)y \quad iff \quad (\exists i \in I)(x, y \in X_i),$$

is an equivalence relation whose equivalence classes are exactly the blocks X_i.

Putting Propositions 4 and 5 together we obtain the useful fact there is a bijection between $\mathrm{Equiv}(X)$ and $\mathrm{Part}(X)$.

Therefore, in principle, it is a matter of taste whether we prefer to work with equivalence relations or partitions.

In computer science, it is often preferable to work with partitions, but not always.

Proposition 6: *Given any set, X, the functions* $\Pi\colon \mathrm{Equiv}(X) \to \mathrm{Part}(X)$ *and* $R : \mathrm{Part}(X) \to \mathrm{Equiv}(X)$ *are mutual inverses, that is,*

$$R \circ \Pi = \mathrm{id} \quad \textit{and} \quad \Pi \circ R = \mathrm{id}.$$

Consequently, there is a bijection between the set, $\mathrm{Equiv}(X)$, *of equivalence relations on X and the set,* $\mathrm{Part}(X)$, *of partitions on X.*

If $f : X \to Y$ is a surjective function, we have the equivalence relation, $\equiv_f$, defined by

$$x \equiv_f y \qquad \text{iff} \quad f(x) = f(y).$$

Then, the equivalence class of any $x \in X$ is the inverse image, $f^{-1}(f(x))$, of $f(x) \in Y$.

Therefore, there is a bijection between $X/\equiv_f$ and Y. Thus, we can identify f and the projection, π, from X onto $X/\equiv_f$.

If f is not surjective, note that f is surjective onto $f(X)$ and so, we see that f can be written as the composition

$$f = i \circ \pi,$$

where $\pi : X \rightarrow f(X)$ is the canonical projection and $i : f(X) \rightarrow Y$ is the *inclusion function* mapping $f(X)$ into Y (*i.e.*, $i(y) = y$, for every $y \in f(X)$).

Given a set, X, the inclusion ordering on $X \times X$ defines an ordering on binary relations on X, namely,

$$R \leq S \quad \text{iff} \quad (\forall x, y \in X)(xRy \Rightarrow xSy).$$

When $R \leq S$, we say that *R refines S*.

If R and S are equivalence relations and $R \leq S$, we observe that every equivalence class of R is contained in some equivalence class of S.

Actually, in view of Proposition 4, we see that *every equivalence class of S is the union of equivalence classes of R*.

We also note that id_X is the least equivalence relation on X and $X \times X$ is the largest equivalence relation on X.

This suggests the following question: Is $\mathrm{Equiv}(X)$ a lattice under refinement?

The answer is yes. It is easy to see that the meet of two equivalence relations is $R \cap S$, their intersection.

But beware, their join is not $R \cup S$, because in general, $R \cup S$ is not transitive.

However, there is a least equivalence relation containing R and S, and this is the join of R and S. This leads us to look at various closure properties of relations.

Transitive Closure, Reflexive and Transitive Closure, Smallest Equivalence Relation

Let R be any relation on a set X. Note that R is reflexive iff $\mathrm{id}_X \subseteq R$. Consequently, the smallest reflexive relation containing R is $\mathrm{id}_X \cup R$. This relation is called the *reflexive closure of R*.

Note that R is transitive iff $R \circ R \subseteq R$, where $xR \circ Ry$ *iff* $(\exists z)\ xRz,\ zRy$. This suggests a way of making the smallest transitive relation containing R (if R is not already transitive). Define R^n by induction as follows:

$$\begin{aligned} R^0 &= \mathrm{id}_X \\ R^{n+1} &= R^n \circ R. \end{aligned}$$

Definition 7: Given any relation, R, on a set, X, the *transitive closure of R* is the relation, R^+, given by

$$R^+ = \bigcup_{n \geq 1} R^n.$$

The *reflexive and transitive closure of R* is the relation, R^*, given by

$$R^* = \bigcup_{n \geq 0} R^n = \mathrm{id}_X \cup R^+.$$

Proposition 7: *Given any relation, R, on a set, X, the relation R^+ is the smallest transitive relation containing R and R^* is the smallest reflexive and transtive relation containing R.*

If R is reflexive, then it is easy to see that $R \subseteq R^2$ and so, $R^k \subseteq R^{k+1}$ for all $k \geq 0$.

From this, we can show that if X is a finite set, then there is a smallest k so that $R^k = R^{k+1}$.

In this case, R^k is the reflexive and transitive closure of R. If X has n elements it can be shown that $k \leq n - 1$.

Note that a relation, R, is symmetric iff $R^{-1} = R$, where $x\, R^{-1}\, y$ iff y R x.

As a consequence, $R \cup R^{-1}$ is the smallest symmetric relation containing R.

This relation is called the *symmetric closure of R*.

Finally, given a relation, R, what is the smallest equivalence relation containing R? The answer is given by

Proposition 8: *For any relation, R, on a set, X, the relation*

$$(R \cup R^{-1})^*$$

is the smalest equivalence relation containing R.