Unit EO

Equivalence and Order

Section 1: Equivalence

The concept of an *equivalence relation* on a set is an important descriptive tool in mathematics and computer science. It is not a new concept to us, as "equivalence relation" turns out to be just another name for "partition of a set." Our emphasis in this section will be slightly different from our previous discussions of partitions in Unit SF. In particular, we shall focus on the basic conditions that a binary relation on a set must satisfy in order to define a partition. This "local" point of view regarding partitions is very helpful in many problems. We start with the definition.

Definition 1 (Equivalence relation) *An equivalence relation on a set S is a partition $\mathcal{K}$ of S. We say that $s, t \in S$ are equivalent if and only if they belong to the same block of the partition $\mathcal{K}$. We call a block an equivalence class of the equivalence relation.*

If the symbol $\equiv$ denotes the equivalence relation, then we write $s \equiv t$ to indicate that s and t are equivalent (in the same block) and $s \not\equiv t$ to denote that they are not equivalent.

Here's a trivial equivalence relation that you use all the time. Let S be any set and let all the blocks of the partition have one element. Two elements of S are equivalent if and only if they are the same. This rather trivial equivalence relation is, of course, denoted by "=".

Example 1 (All the equivalence relations on a set) Let $S = \{a, b, c\}$. What are the possible equivalence relations on S? Every partition of S corresponds to an equivalence relation, so listing the partitions also lists the equivalence relations. Here they are with the equivalences other than $a \equiv a$, $b \equiv b$ and $c \equiv c$, which are always present.

$\big\{\{a\}, \{b\}, \{c\}\big\}$	no others
$\big\{\{a\}, \{b, c\}\big\}$	$b \equiv c$, $c \equiv b$
$\big\{\{b\}, \{a, c\}\big\}$	$a \equiv c$, $c \equiv a$
$\big\{\{c\}, \{a, b\}\big\}$	$a \equiv b$, $b \equiv a$
$\big\{\{a, b, c\}\big\}$	$a \equiv b$, $b \equiv a$, $a \equiv c$, $c \equiv a$, $b \equiv c$, $c \equiv b$

What about the set $\{a, b, c, d\}$? There are 15 equivalence relations. For a five element set there are 52. As you can see, the number increases rapidly. □

Example 2 (Classification by remainder) Now let the set be the integers $\mathbb{Z}$. Let's try to define an equivalence relation by saying that n and k are equivalent if and only if they have the same remainder when divided by 24. In other words, n and k differ by a multiple of 24. Is this an equivalence relation? If it is we should be able to find the blocks of the partition. There are 24 of them, which we could number $0, \ldots, 23$. Block j consists of all integers which equal j plus a multiple of 24; that is, they have a remainder of j when divided by 24. We write this block as $24\mathbb{Z} + j$. If two integers a and b are equivalent, we write $a = b \pmod{24}$ or $a \equiv b \pmod{24}$.

For this to be an equivalence relation, we must verify that the 24 sets $24\mathbb{Z}+j$, $0 \le j \le 23$, are a partition of $\mathbb{Z}$. Each of these blocks is clearly nonempty. Their union is $\mathbb{Z}$ since every integer has some remainder $0 \le j \le 23$ when divided by 24. Why are they disjoint? If $x \in 24\mathbb{Z} + j$ and $x \in 24\mathbb{Z} + k$ where $j \neq k$, then x would have two different remainders j and k when divided by 24. This is impossible.

Of course there is nothing magic about 24 — any positive integer would work. We studied this in Section 1 of Unit NT. We proved an interesting property of these equivalence there: They can be added, subtracted and multiplied. In other words, if $a \equiv A$ and $b \equiv B$, then $a + b \equiv A + B$, $a - b \equiv A - B$ and $ab \equiv AB$. □

Example 3 (Using coimage to define equivalence) Suppose S and T are sets and $F : S \to T$ is a function. We can use the coimage of F to define an equivalence relation. Recall from Section 2 of Unit SF that the coimage of F is the partition of S given by

$$\text{Coimage(F)} = \{F^{-1}(b) \mid b \in T,\ F^{-1}(b) \neq \emptyset\} = \{F^{-1}(b) \mid b \in \text{Image}(F)\}.$$

Since the coimage is a partition of S, it defines an equivalence relation. Thus $s_1 \equiv s_2$ if and only if $F(s_1) = F(s_2)$. As you can see, the idea is to define $F(s)$ to be the property of s that we are interested in. Almost all our examples are of this kind.

In the previous example, we could take $S = \mathbb{Z}$ and define $F(n)$ to be the remainder when n is divided by 24.

The sets need not be numeric. For example, let S be the set of people and let $T = S \times S$. Define

$$F(s) = (\text{mother of } s, \text{ father of } s).$$

Then $s_1 \equiv s_2$ if and only if s_1 and s_2 are siblings (i.e., have the same parents). □

Example 4 (Some equivalence classes of functions) We use the notation $\underline{k}$ to denote $\{1, 2, \ldots, k\}$, the set of the first k positive integers. Consider all functions $S = \underline{m}^{\underline{n}}$. Here are some partitions based on the fact that S is a set of functions:

- We could partition the functions f into blocks according to the sum of the integers in the Image(f). In other words, $f \equiv g$ means $\sum_{i=1}^n f(i) = \sum_{i=1}^n g(i)$.
- We could partition the functions f into blocks according to the max of the integers in Image(f). In other words, $f \equiv g$ means that the maximum values of f and g are the same.
- We could partition the functions f into blocks according to the vector $v(f)$ where the i^{th} component $v_i(f)$ is the number of times f takes the value i; that is, $v_i(f) = |f^{-1}(i)|$.

For example, suppose $f, g, h \in \underline{3}^{\underline{4}}$ in one-line notation are $f = (3, 1, 2, 1)$, $g = (1, 1, 3, 2)$ and $h = (1, 2, 3, 2)$. Then $f \equiv g \not\equiv h$ because $v(f) = v(g) = (2, 1, 1)$ but $v(h) = (1, 2, 1)$.

Each of these defines a partition of S and hence and equivalence relation on S.

We can use the coimage idea of the previous example to describe these equivalence relations:

- $F(f) = \sum_{i=1}^{n} f(i)$,
- $F(f) = \max(f)$ and
- $F(f) = v(f)$. □

Definition 2 (Binary relation on a set) *Given a set S, a binary relation on S is a subset R of $S \times S$. Given a binary relation R, we will write $s\,R\,t$ if and only if $(s, t) \in R$.*[1]

Example 5 (Equivalence relations as binary relations) Suppose $\equiv$ is an equivalence relation on S associated with the partition $\mathcal{K}$. Then the set $R = \{(s, t) \mid s \equiv t\} \subseteq S \times S$ is a binary relation on S associated with the equivalence relation. Thus an equivalence relation is a binary relation.

The converse need not be true. For example $x\,R\,y$ if and only if $x < y$ defines a binary relation on $\mathbb{Z}$, but it is not an equivalence relation because we never have $x < x$, but an equivalence relation requires $x \equiv x$ for all x. □

When is a binary relation an equivalence relation? The next theorem provides necessary and sufficient conditions for a binary relation to be an equivalence relation. Verifying the conditions is a sometimes a useful way to prove that some particular situation is an equivalence relation.

Theorem 1 (Reflexive, symmetric, transitive) *Let S be a set and suppose that we have a binary relation R on S. This binary relation is an equivalence relation if and only if the following three conditions hold.*

(i) (Reflexive) For all $s \in S$ we have $s\,R\,s$.

(ii) (Symmetric) For all $s, t \in S$ such that $s\,R\,t$ we have $t\,R\,s$.

(iii) (Transitive) For all $r, s, t \in S$ such that $r\,R\,s$ and $s\,R\,t$ we have $r\,R\,t$.

Proof: We first prove that an equivalence relation satisfies (i)–(iii). Suppose that $\equiv$ is an equivalence relation. Since s belongs to whatever block it is in, we have $s \equiv s$. Since $s \equiv t$ means that s and t belong to the same block, we have $s \equiv t$ if and only if we have $t \equiv s$. Now suppose that $r \equiv s$ and $s \equiv t$. Then r and s are in the same block and s and t are in the same block. Thus r and t are in the same block and so $r \equiv t$.

[1] A binary relation is a special case of a relation from S to T (discussed in Unit SF), namely, $T = S$.

We now suppose that (i)–(iii) hold and prove that we have an equivalence relation; that is, a partition of the set S. What would the blocks of the partition be? Everything equivalent to a given element should be in the same block. Thus, for each $s \in S$ let $B(s)$ be the set of all $t \in S$ such that $s\ R\ t$. We must show that the set of these sets form a partition of S; that is, $\{B(s) \mid s \in S\}$ is a partition of S.

In order to have a partition of S, we must have

(a) the $B(s)$ are nonempty and every $t \in S$ is in some $B(s)$ and

(b) for every $p, q \in S$, $B(p)$ and $B(q)$ are either equal or disjoint.

Since R is reflexive, $s \in B(s)$, proving (a). We now turn our attention to (b). Suppose $B(p) \cap B(q)$ is not empty. We must prove that $B(p) = B(q)$. Suppose $x \in B(p) \cap B(q)$ and $y \in B(p)$. We have, $p\ R\ x$, $q\ R\ x$ and $p\ R\ y$. By the symmetric law, $x\ R\ p$. Using transitivity twice:

$$\begin{aligned}
\big(q\,R\,x \;\text{ and }\; x\,R\,p\big) &\;\text{ implies }\; q\,R\,p,\\
\big(q\,R\,p \;\text{ and }\; p\,R\,y\big) &\;\text{ implies }\; q\,R\,y.
\end{aligned}$$

By the definition of B, this means $y \in B(q)$. Since this is true for all $y \in B(p)$, we have proved that $B(p) \subseteq B(q)$. Similarly $B(q) \subseteq B(p)$ and so $B(p) = B(q)$. This proves (b). □

Example 6 (The rational numbers) Let the set be $S = \mathbb{Z} \times \mathbb{Z}^*$, where $\mathbb{Z}^*$ is the set of all integers except 0. Write $(a, b) \equiv (c, d)$ if and only if $ad = bc$.

We now use Theorem 1 to prove that, in fact, this is an equivalence relation.

(i) To verify $(a, b) \equiv (a, b)$, we must check that $ab = ba$ (since $c = a$ and $d = b$). This is obviously true.

(ii) Suppose $(a, b) \equiv (c, d)$. This means $ad = bc$. To verify that $(c, d) \equiv (a, b)$, we must check that $cb = da$. This follows easily from $ad = bc$.

(iii) Suppose $(a, b) \equiv (c, d)$ and $(c, d) \equiv (e, f)$. We must verify that $(a, b) \equiv (e, f)$. In other words, we are given $ad = bc$ and $cf = de$, and we want to conclude that $af = be$. We have $(ad)(cf) = (bc)(de)$. We're done if we can cancel c and d from both sides of this equality. We can cancel d since $d \in \mathbb{Z}^*$. If $c \neq 0$ we can cancel it, too, and we're done. What if $c = 0$? In that case, since $d \neq 0$, it follows from $ad = bc$ that $a = 0$. Similarly, it follows from $cf = de$ that $e = 0$. Thus $af = be = 0$.

With a moment's reflection, you should see that $ad = bc$ is a way to check if the two fractions a/b and c/d are equal. We can label each equivalence class with the fraction a/b that it represents.

Just because we've defined fractions, that doesn't mean we can do arithmetic with them. We need to prove that equivalence classes can be added, subtracted, multiplied, and divided. This problem was mentioned for modular arithmetic at the end of Example 2. We won't prove it for rational numbers. □

In the exercises, you will acquire more practice with equivalence relations. You should find them easier to deal with than the previous two examples.

The Pigeonhole Principle

We now look at a class of problems that relate to various types of restrictions on equivalence relations. These problems are a part of a much more general and often very difficult branch of mathematics called *extremal set theory*. The following theorem is a triviality, but its name and some of its applications are interesting. You should be able to prove the theorem.

Theorem 2 (Pigeonhole principle) *Suppose $\mathcal{K}$ is a partition of a set S and $|S| = s$. If $\mathcal{K}$ has fewer than s blocks, then some block must have at least two elements.*

Where did the name come from? Old style desks often had an array of small horizontal boxes for storing various sorts of papers — unpaid bills, letters, etc. These were called pigeonholes because they often resembled the nesting boxes in pigeon coops. Imagine slips of paper, with one element of S written on each slip. Put the slips into the boxes. At least one box must receive more than one slip if there are more slips than boxes — that's the pigeonhole principle. After the slips are in the boxes, a partition of the set of slips has been defined. (The boxes are the blocks.)

Example 7 (Applying the pigeonhole principle) Designate the months of the year by the set numbers $M = \underline{12} = \{1, 2, \ldots, 12\}$. What is the smallest integer k such that among any k people, there must be at least two people with the same first letter of their last name and same birth month? This is a typical application of the pigeonhole principle.

Recall that we can define a partition of a set P by defining a function f with domain P and letting the partition of P be the coimage of f. In this case, we let P be the set of people and define $f : P \to M \times A$, where A is the set of letters in the alphabet, as follows: $f(p) = (m, a)$, where m is the month in which p was born and a is the first letter of p's last name. To be able to apply the pigeonhole principle to obtain the conclusion asked for, we must have $|M \times A| < k$, where $k = |P|$. In other words, we must have $12 \times 26 = 312 < k$. The smallest such $k = 313$. If $k = 312$ then it is possible to have a group of people, no two of which have the same first letter of their last name and same birth month.

What we did in the last example is a common way to set things up for the pigeonhole principle. We are given a set S (people) and a property (birthdays and initials) of the elements of S. We want to know that two elements of S have equal properties. To use the coimage, define a function $f : S \to T$ where T is the set of possible properties. Since the coimage can have at most $|T|$ blocks, we need $|T| < |S|$ to use the pigeonhole principle.

Example 8 (A divisibility example) Here is a somewhat less trivial example. Let $n > 0$ be a an integer. What is the smallest integer k such that given any set $\{t_1, t_2, \ldots, t_k\}$ of distinct integers, either there must be some $i \neq j$ such that $n \mid (t_i - t_j)$, or there must be some i such that $n \mid t_i$? We must give a little thought as to how to use the coimage version of the pigeonhole principle in this situation. It seems that the correct function is $f : \{t_1, t_2, \ldots, t_k\} \to \{0, 1, \ldots, n-1\}$ where $f(x) = x \pmod{n}$. Suppose first that there is no i such that $n \mid t_i$. Then $|\text{Image}(f)| < n$ because the only values that can be in the image are $1, 2, \ldots, (n-1)$. In that case, if $k \geq n$ then we can apply the pigeonhole principle. There must be at least one pair, $i < j$, such that $f(t_i) = f(t_j)$ or, equivalently, $t_i \pmod{n} = t_j \pmod{n}$. For this pair, $n \mid (t_i - t_j)$. If $k = n-1$, we can take the set $\{1, 2. \ldots, n-1\}$ for which there is neither some i such that $n \,|t_i$ nor a pair $i \neq j$ such that $n \mid (t_i - t_j)$. Thus, the smallest k is $k = n$. You should think carefully about this example as it typical of one type of analysis associated with these extremal problems. The idea is that one condition (in this case $n \mid t_i$ for some i), when negated, brings about a situation where the pigeonhole principle applies.

Results gotten from the application of the pigeonhole principle are sometimes used in clever ways to get new results. Suppose that we have a sequence $a_1, a_2, \ldots, a_n$ of positive integers. Then, by the previous paragraph applied to the "partial sums" of this sequence, there must be a partial sum, $t_i = a_1 + a_2 + \cdots + a_i$ such that either $n \mid t_i$ or there is a pair $i < j$ such that $n \mid (t_j - t_i)$. Note that $t_j - t_i = a_{i+1} + \cdots + a_j$. Thus, if you have any sequence of n positive integers, there is at least one "consecutive" sum of the form $\sum_{k=p}^{q} a_k$ that is divisible by n. For example, take $n = 8$ and take the sequence to be 11, 12, 23, 5, 7, 9, 21, 9. The consecutive sum $12 + 23 + 5 = 40$ is divisible by 8. In general, this consecutive sum that is divisible by n will, as in this example, not be unique.

Example 9 (Equal sums) Given a positive integer N, how large must t be so that, for every list $A = (a_1, \ldots, a_t)$ of t integers, there are $i \neq j$ and $k \neq m$ such that $a_i + a_j = a_k + a_m \pmod{N}$ and $\{i, j\} \neq \{k, m\}$. (The last condition avoids the trivial situation $a_i + a_j = a_i + a_j$.) For example, if $N = 5$ and $A = (1, 2, 5, 8)$, then $1 + 2 = 5 + 8 \pmod{5}$. We want to look at sums of pairs of elements of A, so our set S will be pairs of indices chosen from $\underline{t}$ and $f(i, j) = a_i + a_j \pmod{N}$ defines a function $f : S \to \{0, 1, \ldots, N-1\}$. Thus there are N pigeonholes. What is $|S|$? In other words how many pairs can be chosen from $\underline{t}$? The answer is $\binom{t}{2} = (t^2 - t)/2$. For the pigeonhole principle we need $(t^2 - t)/2 > N$, which is equivalent to $t^2 - t - 2N > 0$. The curve $y = x^2 - x - 2N$ is a parabola opening upward. Thus $y > 0$ whenever x exceeds the larger root of the quadratic. Solving for the larger root and setting t greater than it, we obtain

$$t > \frac{1 + \sqrt{1 + 8N}}{2} \quad \text{implies that} \quad a_i + a_j = a_k + a_m$$

for two distinct pairs of distinct indices $\{i, j\} \neq \{k, m\}$. □

***Example 10 (Subset sums)** In the set $S = \{1, 2, 3, 4\}$ there are two different (not equal as sets) subsets $P \subset S$ and $Q \subset S$ such that the sum of the entries in P (designated by $\sum P$) equals the sum of the entries in Q (take $P = \{1, 2, 3\}$ and $Q = \{2, 4\}$). We say that the set $S = \{1, 2, 3, 4\}$ has the *two-sum property.* In the set $S = \{1, 2, 4, 8\}$ there is no such pair of subsets. This set fails to have the two-sum property. Suppose we start with a set

$U = \underline{n} = \{1, 2, \ldots, n\}$. What conditions can we put on k such that *every* subset $S \subset U$ of size k has the two-sum property; i.e., S has a pair of distinct subsets P and Q with $\sum P = \sum Q$? This question has the aura of a pigeonhole problem, but how do we describe the function and coimage?

We will look just at the case $n = 16$. Thus $U = \underline{16}$. What is the largest subset of U that does *not* have the two-sum property? Call the size of this set $T(16)$. Using the pigeonhole principle, we will show that every $S \subset U$ of size 7 has the two-sum property. You can check that $\{1, 2, 4, 8, 16\}$, a subset of size 5, does not have the two-sum property.[2] Thus $T(16) \geq 5$. Assuming the above mentioned result for sets of size 7, we now know that either $T(16) = 5$ or $T(16) = 6$. After searching for a subset of size 6 that does not have the two-sum property, you may become convinced that $T(16) = 5$. Unless you looked at all of the $\binom{16}{6} = 8,008$ subsets, this is not a proof. Using a computer program to do this, we found that each of these 8,008 subsets has the two-sum property and so $T(16) = 5$.

Such gaps (Is $T(16) = 5$ or 6?) are common in applications of the pigeonhole principle. Any careful study of this problem would have to go into such gaps between the largest counterexample we have found ($k = 5$) and the smallest k for which the pigeonhole principle works ($k = 7$). This is an annoying feature of the pigeonhole principle in many cases where it is applied.

We now show that every $S \subseteq \underline{16}$ with $|S| = 7$ has the two-sum property.

A subset $S \subseteq U$ with $|S| = k$ has 2^k subsets. If we are going to find distinct subsets P and Q of S with $\sum P = \sum Q$, then clearly neither P nor Q can be empty or equal to S. It seems that we want to apply the pigeonhole principle to the set K of all subsets of a set S, except $\emptyset$ and S. There are $2^k - 2$ such subsets. The function $f : K \to R$ will be given by $f(P) = \sum P$ for $P \in K$. What is R? It is all possible subset sums. We need to work this out.

The value of the sum of the entries over such a subset can be as small as 1 and as large as $r = (16 - (k-2)) + (16 - (k-3)) + \cdots + 16$ (the largest sum of any $k - 1$ elements from U). The pigeonhole principle assures us that there will be two distinct elements X and Y in K with $f(X) = f(Y)$ if $|K| > |R|$. In other words, if $2^k - 2 > r$. Using a calculator, we see that the first k that satisfies this inequality is $k = 7$. Thus, any subset $S \subseteq \underline{16}$ of size 7 or larger has the two-sum property. □

The pigeonhole principle answers the question

When must a partition of a set have a block of size at least two?

Sometimes that's not enough. For example, what if we want more than two people to have the same birthday? This is as easy to answer as the original question:

Theorem 3 (Extended pigeonhole principle) *Suppose $f : S \to R$ is a function, $|R| = r$ and $|S| = s$. The coimage of f must have a block of size $\lceil s/r \rceil$ or larger.*

[2] You should be able to see why this is so by thinking in terms of the binary number system: every positive integer has a *unique* representation as sums of powers of 2.

Proof: Suppose the Coimage(f) has t blocks. Since $|R| = r$, we have $t \leq r$. Let $\{B_1, B_2, \ldots, B_t\}$ be the blocks of Coimage(f) and suppose $|B_i| < s/r$ for $i = 1, \ldots, t$. Now

$$s = |S| = |B_1| + |B_2| + \cdots + |B_t| < (s/r) + (s/r) + \cdots + (s/r) = (s/r)t \leq (s/r)r = s,$$

a contradiction. Thus some B_i has $|B_i| \geq s/r$. Since the size of a block is an integer and s/r may not be an integer, we can assert that Coimage(f) must have at least one block of size *greater than or equal* to $\lceil s/r \rceil$. □

For example, how many people must we have to be sure that at least k have the same birthday? Let s be the number of people. Since there are 366 possible birthdays, $r = 366$. By the theorem, it suffices to have $\lceil s/366 \rceil \geq k$. In other words, $s/366 > k - 1$. Thus $s > 366(k - 1)$ guarantees that at least k out of s people *must* have the same birthday.

In some instances, the most difficult part of applying the pigeonhole principle is finding the right partition (equivalence relation). The next example is of this type.

***Example 11 (Monotone subsequences)** A sequence is *monotonic* if it is either decreasing or increasing.

Consider the sequence 7, 5, 2, 6, 8, 1, 9. Starting at a term, say 2, we can move to the right selecting a subsequence (not necessarily consecutive) that is increasing: 2, 6, 8, 9. Or, starting at 2 we could do the same, only selecting a decreasing subsequence: 2, 1.

In a general sequence, we cannot guarantee a long increasing subsequence — the numbers in the sequence might be decreasing. Similarly, we cannot guarantee a long decreasing subsequence. However, we might hope to guarantee a long *monotonic* subsequence. For example, the longest decreasing subsequence in our example is 2, 1, but there is a monotonic subsequence of length 4.

Let's use the pigeonhole principle for the general case. Suppose we start with a sequence of length m, $a_1, a_2, \ldots, a_m$, where the terms are distinct real numbers. How long an increasing or decreasing subsequence *must* be present?

Suppose $a_t = a_{d_1}, a_{d_2}, \ldots, a_{d_k}$ is a decreasing subsequence of length k starting at a_t and that $a_t = a_{i_1}, a_{i_2}, \ldots, a_{i_l}$ is an increasing subsequence of length l starting at a_t.

- If $a_{t-1} > a_t$, then $a_{t-1}, a_{d_1}, a_{d_2}, \ldots, a_{d_k}$ is a decreasing subsequence of length $k + 1$ starting at $t - 1$.
- If $a_{t-1} < a_t$, then $a_{t-1}, a_{i_1}, a_{i_2}, \ldots, a_{i_l}$ is an increasing subsequence of length $l + 1$ starting at $t - 1$.

Thus the length of either the increasing or decreasing sequence has increased by 1.

Let's formalize this a bit. Let D_t and I_t be the lengths of the longest decreasing and longest increasing subsequences starting at a_t. We've just shown that either $D_{t-1} = D_t + 1$ or $I_{t-1} = I_t + 1$. Thus $(I_{t-1}, D_{t-1}) \neq (I_t, D_t)$. All we used was that $t-1$ is less than t. Thus we can replace $t-1$ by any $u < t$ in the above argument to conclude that $(I_u, D_u) \neq (I_t, D_t)$ whenever $u < t$. In other words, $f : \underline{m} \to \mathbb{Z} \times \mathbb{Z}$ given by $f(t) = (I_t, D_t)$ is an injection.

We're ready to apply the pigeonhole principle. Why? The pigeonhole principle tells us when f *cannot* be an injection because the pigeonhole principle guarantees that Coimage(f) has a block of size greater than one.

Suppose that the overall longest increasing sequence has length ι ("iota") and the overall longest decreasing sequence has length δ ("delta"). Thus the image of f is contained in $\underline{\iota} \times \underline{\delta}$ and so $|\text{Coimage}(f)| \leq \iota\delta$. Suppose $\iota\delta < m$. By the pigeonhole principle, there must be $p < q$ with $f(p) = f(q)$, a contradiction because we know f is an injection.

We have shown that, if the longest decreasing subsequence has length δ and the longest increasing subsequence has length ι, then the sequence has length at most $\iota\delta$. In other words, $m \leq \iota\delta$.

Put another way, we have shown that, if $m > \alpha\beta$ for some integers α and β, then there must be either an increasing subsequence longer than α or a decreasing subsequence longer than β. For example, if $m = 100$, there is either an increasing subsequence of length at least 12 or a decreasing subsequence of length at least 10 because $100 > 99 = 11 \times 9$. If we write $99 = 9 \times 11$, we see that there is either an increasing subsequence of length at least 10 or a decreasing subsequence of length at least 12.

How long a monotonic subsequence must a sequence of length m have? In this case, we take $\alpha = \beta = n$ because we want to make sure the subsequence is long whether it is increasing or decreasing. If $m > n^2$, there must be a monotonic subsequence of length $n+1$ or greater. Of course, if there is a monotonic subsequence of length at least $n + 1$, then there is one of length exactly $n + 1$: just throw some elements if the subsequence is too long. ☐

Sometimes, in working problems concerning partitions, there are restrictions on block sizes. In such cases, can be useful to list the possible *type vectors*. A type vector $\vec{v}$ for a partition has v_i equal to the number of blocks containing exactly i elements. For example, we list all type vectors for partitions of $S = \underline{15}$ with maximum block size 4 and exactly 5 blocks. Since $v_i = 0$ for $i > 4$, we give just $v_1, \ldots, v_4$.

1	**2**	**3**	**4**
1	1	0	3
1	0	2	2
0	2	1	2
0	1	3	1
0	0	5	0

The top row represents the block sizes, ranging from 1 to maximum size of 4. The remaining rows tell us how many blocks of each size there are (i.e., these rows correspond to the type vectors). You should think carefully about why this list of type vectors is complete for the restrictions given.

In the exercises for this section you are asked to solve a problem related to the following situation: "Fifteen clients are being defended against lawsuits by a group of five lawyers. Each client is assigned exactly one lawyer and no lawyer is to represent less than one or more than four clients." To see the connection between this problem and partitions, note first that the condition that each client be assigned one lawyer specifies that the correspondence between clients and lawyers is a function, with domain size 15 (the clients) and codomain size 5 (the lawyers). The condition that no lawyer is to represent less than one or more than four clients specifies that the function is onto and that the maximum block size of the coimage is four. That the function is onto and the codomain has five elements says that there are exactly five blocks in the coimage. Thus, the above table lists the type vectors of all coimages for allowable correspondences between clients and lawyers.

Exercises for Section 1

1.1. Let S be the set of students in a college. Define students $x \equiv y$ to be related if they have both the same age and the same number of years completed in college. Show that $\equiv$ is an equivalence relation by defining a function a whose coimage is the set of equivalence classes of $\equiv$.

1.2. Let $\mathbb{Z}$ be the integers with d a positive integer. Define integers $x \equiv y$ to be related if $d \mid (x-y)$. Show that $\equiv$ is an equivalence relation by defining a function m whose coimage is the set of equivalence classes of $\equiv$.

1.3. Let $\mathbf{F}_n$ be the set of all statement forms in n Boolean variables. Define forms $x \equiv y$ to be related if they have the same truth table. Show that $\equiv$ is an equivalence relation by defining a function t with coimage the equivalence classes of $\equiv$.

1.4. Let $\mathbb{Z}$ be the integers with d and k positive integers. Define $x \equiv y$ to be related if $d \mid (x^k - y^k)$. Show that $\equiv$ is an equivalence relation on $\mathbb{Z}$ by defining a function m whose coimage is the set of equivalence classes of $\equiv$.

1.5. Let $\mathbb{R}$ be the real numbers. Define $x \equiv y$ to be related if $x - y \in \mathbb{Z}$. Show that $\equiv$ is an equivalence relation on $\mathbb{R}$ by defining a function u whose coimage is the set of equivalence classes of $\equiv$.

1.6. In any group of 677 people, there must be at least two who have the same first and last letters of their names. Explain.

1.7. In each case give an explanation in terms of functions and coimages.

(a) Must any set of $k > 1$ integers have at least two with the same remainder when divided by $k - 1$?

(b) Must any set of $k > 1$ integers have at least two with the same remainder when divided by k?

1.8. What is the smallest integer k such that every k-element subset of the set $S = \underline{n}$ must always contain a pair of elements whose sum is $n + 1$?

1.9. Let $n \geq 1$ be an integer. What is the smallest integer k such that every k-element subset of the set $S = \{0, 1, 2, \ldots, n\}$ must always contain an even integer? Must always contain an odd integer?

1.10. What is the smallest integer k such that any set S of k integers selected from the set $\underline{50} = \{1, 2, \ldots, 50\}$ will always have two distinct integers, $x \in S$ and $y \in S$ such that $\gcd(x, y) > 1$? ($\gcd(x, y)$ is the greatest common divisor of x and y.)

1.11. What is the smallest integer k such that any set S of k people must have at least three people who were born in the same month of the year?

1.12. Let P be a group of 30 people. Let f be the function from P to M, where $M = \underline{12}$ represents the 12 months of the year, and $f(x)$ is the birth-month of x. Among such a group, there need not be any group of four people that have the same birth-month. One way this can happen is if Coimage(f) has ten blocks of size three. Describe the structure of the coimage of all other examples. What are their type vectors?

1.13. Some cultures divide a day into "quarter days" in order to pay respect to the tidal cycle (four six hour tidal cycles in each 24 hour period). There are 1461 ATCs (Annual Tidal Cycles) per solar year. What is the smallest integer k such that among k people there are at least four born in the same ATC?

1.14. There are N students in a class. Their exam scores ranged between 27 and 94. All possible scores were achieved by at least one student except for the scores 31, 43, and 55 (none of the students got these scores). What is the smallest value of N that guarantees that at least three students achieved the same score?

1.15. There are twelve 1967 pennies, seven 1968 pennies, and four 1971 pennies in a jar. Let N_k denote the smallest number of pennies you need to select to guarantee that you have k pennies of the same date. Find N_4, N_6 and N_8.

1.16. Let $t_1, t_2, \ldots t_n$ be n integers. Show that either $n \mid t_k$ for some k or $n \mid (t_i - t_j)$ for some $i \neq j$.

1.17. Let $n > 1$ be an integer. What is the smallest value of k such that, given any k distinct integers, $t_1, t_2, \ldots t_k$, there must be two of them t_i and t_j, $i \neq j$, such that either $n \mid (t_i - t_j)$ or $n \mid (t_i + t_j)$?
Hint: We want t_i and t_j to go in the same pigeonhole if either $t_i = -t_j \bmod n$ (so that $n \mid (t_i + t_j)$) or $t_i = t_j \bmod n$ (so that $n \mid (t_i - t_j)$).

1.18. Let $n > 1$ be an integer. What is the smallest value of m such that, given any m distinct integers, $t_1 < t_2 < \cdots < t_m$, chosen from the set $S = \underline{n}$, there must be

$i < j$, such that $t_i \mid t_j$.
Hint: Remove all factors of 2 from the elements of S.

***1.19.** We want to show that $m \leq \iota\delta$ in Example 11 is best possible. In other words, there exist sequences of length $\iota\delta$ with longest increasing subsequence of length ι and longest decreasing subsequence of length δ.

(a) Construct a sequence of length ι whose longest decreasing subsequence has length $\delta = 1$ and whose longest increasing subsequence has length ι.

(b) Construct a sequence of length 2ι whose longest decreasing subsequence has length $\delta = 2$ and whose longest increasing subsequence has length ι.

(c) Construct a sequence of length $\delta\iota$ whose longest decreasing subsequence has length δ and whose longest increasing subsequence has length ι.

***1.20.** Suppose $m = pq$ and an m-long sequence of distinct real numbers does not have a p-long decreasing subsequence. Prove that it has has a q-long increasing subsequence.

***1.21.** Suppose $m > n^4$. Let (a_1, b_1), ..., (a_m, b_m) be an m-long sequence where the a_i and b_j are distinct real numbers. The goal of this exercise is to prove that there is an $(n+1)$-long subsequence $(a_{t_1}, b_{t_1}), \ldots, (a_{t_{n+1}}, b_{t_{n+1}})$ such that the sequences $a_{t_1}, \ldots, a_{t_{n+1}}$ and $b_{t_1}, \ldots, b_{t_{n+1}}$ are both monotone.

(a) Let $k = n^2 + 1$. Prove that the sequence $a_1, \ldots, a_m$ has a k-long monotone subsequence. Call it $a_{s_1}, \ldots, a_{s_k}$.

(b) Prove that the subsequence $b_{s_1}, \ldots, b_{s_k}$ has an $(n+1)$-long monotone subsequence. Call it $b_{t_1}, \ldots, b_{t_{n+1}}$.

(c) Prove that the indices $t_1, \ldots, t_{n+1}$ solve the problem.

1.22. Fifteen clients are being defended against lawsuits by a group of five lawyers. Each client is assigned exactly one lawyer and no lawyer is to represent less than one or more than four clients. Show that if two lawyers are assigned less than three clients, at least two must be assigned four clients.

Section 2: Order

In Theorem 1 we showed the connection between certain binary relations on a set S and partitions of the same set. In this section we will study binary relations that are, as before, reflexive and transitive, but, instead of being symmetric, are "antisymmetric." We begin by defining the most general idea of a relation from one set to another and, specializing from that, defining the central theme of this section, order relations.

Definition 3 (Order relation, partially ordered set, poset) *Binary relations are defined in Definition 2. If R is a binary relation on S, then $(x, y) \in R$ is also denoted by $x\,R\,y$. Likewise, $(x, y) \notin R$ is denoted by $x \not R\, y$.*

A binary relation on a set S is called an order relation if it satisfies the following three conditions and then it is usually written $x \preceq y$ instead of $x\ R\ y$.

(i) (Reflexive) For all $s \in S$ we have $s \preceq s$.

(ii) (Antisymmetric) For all $s, t \in S$ such that $s \neq t$, if $s \preceq t$ then $t \not\preceq s$.

(iii) (Transitive) For all $r, s, t \in S$ such that $r \preceq s$ and $s \preceq t$ we have $r \preceq t$.

A set S together with an order relation $\preceq$ is called a partially ordered set or poset. Formally, a poset is a pair $(S, \preceq)$. We shall, once the binary relation is defined, refer to the poset by the set S alone, not the pair.

Order relations and equivalence relations appear similar: Both are reflexive and transitive. The only difference is that one is antisymmetric and the other is symmetric. Although this may seem a small difference, it makes a *big* difference in two types of relations, as you'll see if you compare the examples of order relations in this section with the examples of equivalence relations in the previous section.

If we use the alternative notation R for the relation, then the three conditions for an order relation are written as follows.

(i) For all $s \in S$ we have $(s, s) \in R$.

(ii) For all $s, t \in S$ such that $s \neq t$, if $(s, t) \in R$ then $(t, s) \notin R$.

(iii) For all $r, s, t \in S$ such that $(r, s) \in R$ and $(s, t) \in R$ we have $(r, t) \in R$.

In the exercises for this section, you will get a chance to think about various binary relations on a set S that may or may not satisfy the conditions of being reflexive, symmetric, antisymmetric, or transitive.

Example 12 (Some partially ordered sets) You are already familiar with a number of basic sets S with order relations R (*posets* for short).

Total orders: Let $S = \mathbb{Z}$, the integers, and define $n\,R\,m$ if $n \leq m$ (usual ordering on integers). Clearly $n \leq n$ for all $n \in \mathbb{Z}$ (reflexive condition). For $n, m \in \mathbb{Z}$, if $n \neq m$ and $n \leq m$, then $m \not\leq n$ (antisymmetric condition). For all p, q, and r in $\mathbb{Z}$, if $p \leq q$ and $q \leq r$, then $p \leq r$ (transitive condition). We call the relation $\leq$, the *natural ordering* of the integers. This same ordering applies to any subset of $\mathbb{Z}$. For example,take $S = \underline{n}$ (the first n positive integers) ordered by $\leq$. This S is a poset. The natural ordering on the integers has an additional property, namely, for all $n, m \in \mathbb{Z}$, either $m \leq n$ or $n \leq m$. The order relation $\leq$ is called a *total ordering* or *linear ordering* because, for any two elements x and y, either $x \leq y$ or $y \leq x$ (or both if $x = y$). The relation $\leq$ can be extended to the real numbers $\mathbb{R}$ and the rational numbers $\mathbb{Q}$.

Subset lattice: Given a set X, let $S = \mathcal{P}(X)$ be the power set of X (the set of all subsets of X). For $A, B \in \mathcal{P}(X)$ we can define $A\,R\,B$ by $A \subseteq B$. The relation $\subseteq$ is an order relation (called *set inclusion*). To check the conditions that an order relation must satisfy, note that $A \subseteq A$ for all $A \in \mathcal{P}(X)$. For all $A, B \in \mathcal{P}(X)$, if $A \neq B$ and $A \subseteq B$, then there is some $x \in B$, $x \notin A$. Thus, $B \not\subseteq A$. We leave transitivity for you to check. The poset $\mathcal{P}(X)$ with the relation $\subseteq$ is called the *lattice of subsets* of X.[3] The subset lattice has the

[3] "Lattice" is a technical term whose meaning we will not explain. It is a poset with certain additional properties.

property that, if $|X| > 1$ then there are always elements $A, B \in \mathcal{P}(X)$ such that $A \not\subseteq B$ and $B \not\subseteq A$. For example, if $X = \{a, b\}$, then $A = \{a\}$ and $B = \{b\}$ are such *incomparable* subsets. If $X = \{a, b, c\}$, then all three subsets of size two are pairwise incomparable. If $|X| = n > 1$, then the set $\mathcal{P}_k(X)$ of all subsets of size k, for any $0 < k < n$, is always a nontrivial (at least two elements) collection of pairwise incomparable subsets of X. The number of elements in $\mathcal{P}_k(X)$ is the binomial coefficient $C(n, k)$.

Divides relation: Another familiar poset is gotten by taking a collection of positive integers, say $S = \underline{n}$, and defining the relation $i\ R\ j$ to be $i \mid j$ (the *divides* relation). Clearly, for all $i \in S$, $i \mid i$. For all $i, j \in S$, if $i \neq j$ and $i \mid j$, then j does not divide i (antisymmetry). If $i \mid j$ and $j \mid k$, then $i \mid k$ (transitivity). □

We recall some definitions that were mentioned in the previous example.

Definition 4 (Incomparable elements, linear order, total order, chain) *Let $(S, \preceq)$ be a poset. If $x, y \in S$ and neither $x \preceq y$ nor $y \preceq x$ are true, we call x and y incomparable. If either $x \preceq y$ or $y \preceq x$, we say that x and y are comparable.*

If **every** *two elements of S are comparable, $(S, \preceq)$ is called a linear order. It is also called a total order or a chain.*

Since some students have the most trouble with so called "trivial" situations, we take a look at those in the next example.

Example 13 (Trivial examples of binary relations) Suppose first that $S = \emptyset$. Since S is empty, so is $S \times S$. Hence the only binary relation on S is $R = \emptyset$. Is R reflexive? symmetric? transitive? Yes. One way to see this is to note that the conditions talk about all $s \in S$ (and possibly r and t). Since there is nothing in S, there is nothing to check. Another way to see this is to look at how you show a condition is *not* satisfied. For example, the way you show something is **not** reflexive is to find an $x \in S$ such that $(x, x) \notin R$. Since S is empty, it is impossible to find such an $x \in S$. Thus, R is reflexive. Similarly, it is also symmetric, antisymmetric and transitive.

Suppose S is not empty, but R is empty. Then, if we choose any $x \in S$, $(x, x) \notin R$. Thus, R is not reflexive. What about symmetric? To show that R is *not* symmetric, we need to find $x, y \in S$ such that $(x, y) \in R$, but ... (stop right here). We can't do this because R is empty. Thus R is symmetric. For the same basic reason, R is also antisymmetric and transitive.

Most cases are of the form $R \subseteq S \times S$ with R (and thus S) not empty. The smallest case is $|S| = 1$ and $|R| = 1$. In this case, if $S = \{a\}$, then $S \times S = \{(a, a)\}$ and so $R = \{(a, a)\}$. You should be able to verify that R is reflexive, symmetric, antisymmetric and transitive.

The next simplest case is $|S| = 2$. Things are suddenly more complicated. There are four elements in $S \times S$. Thus $2^4 = 16$ choices for R, fifteen of which are nonempty. To get a feeling for the situation, we look at some *incidence matrices* for R. These are 2×2 matrices whose rows and columns are labeled with the elements of S. The entry (x, y) is 1 if $(x, y) \in R$ and is 0 if $(x, y) \notin R$. Here are six of the fifteen possible incidence matrices for $S = \{a, b\}$:

$$\underset{\mathbf{A}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 1 \\ b & 0 & 0 \\ \hline \end{array}} \quad \underset{\mathbf{B}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 0 \\ b & 0 & 1 \\ \hline \end{array}} \quad \underset{\mathbf{C}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 1 \\ b & 1 & 0 \\ \hline \end{array}} \quad \underset{\mathbf{D}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 0 \\ b & 1 & 0 \\ \hline \end{array}} \quad \underset{\mathbf{E}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 0 \\ b & 1 & 1 \\ \hline \end{array}} \quad \underset{\mathbf{F}}{\begin{array}{c|cc|} & a & b \\ \hline a & 1 & 1 \\ b & 1 & 1 \\ \hline \end{array}}$$

The first matrix **A** describes the binary relation $\{(a, a), (a, b)\}$ because $\mathbf{A}(a, a) = \mathbf{A}(a, b) = 1$ and $\mathbf{A}(b, a) = \mathbf{A}(b, b) = 0$. The entries in positions (a, a) and (b, b) are called the entries on the "main diagonal" of **A**. Positions (a, b) and (b, a) are "symmetric off-diagonal positions" about the main diagonal and their entries are "symmetric off-diagonal entries" with respect to **A**. The relation defined by **A** (or more simply, "the matrix **A**") is not reflexive, is antisymmetric, is not symmetric, and is transitive. With regard to the last statement, note that, since $R = \{(a, a), (a, b)\}$, there is not really anything to check since combining (a, a) with (a, b) using transitivity just gives us (a, b) again. This is always true: we never need to use a diagonal entry like (a, a) in checking transitivity. On the other hand, matrix **C** is *not* transitive because $(b, a) \in R$ and $(a, b) \in R$ would give us $(b, b) \in R$, which is not true. Note that there is already a lesson here. If symmetric off-diagonal entries are both 1, but either of their corresponding diagonal entries is not 1, then the relation is not transitive. This is true for the incidence matrix of any relation. For the various matrices we have

	reflexive	symmetric	antisymmetric	transitive
A	no	no	yes	yes
B	yes	yes	yes	yes
C	no	yes	no	no
D	no	no	yes	yes
E	yes	no	yes	yes
F	yes	yes	no	yes

You should make sure you understand the reasons for all of these statements. □

Example 14 (Counting relations) As we have seen in the previous example, relations on a set S correspond to matrices of zeroes and ones (the incidence matrices). A relation on a four element set, for example, corresponds to a 4×4 matrix of zeroes and ones (an incidence matrix). The matrix below, with appropriate substitutions of 0's and 1's for the symbolic entries, is such an incidence matrix. The rows and columns should be labeled with the four elements of the set (as in the previous example, where we worked with 2×2 incidence matrices), but we omit that here for simplicity.

Since an $n \times n$ matrix has n^2 entries and each entry can be either a 0 or a 1, there are 2^{n^2} such matrices. This number grows *very* rapidly; for example, when $n = 4$ we have $2^{4^2} = 2^{16} = 65,536$.

Let's look at 4×4 matrices We could think of starting with a matrix, such as the one below, labeled with symbols d_i (for diagonal), u_i (for upper), and l_i (for lower), and then, in some manner (just how is up to us) substituting zeroes and ones for the sixteen symbols.

$$\begin{array}{cccc} d_1 & u_1 & u_2 & u_3 \\ l_1 & d_2 & u_4 & u_5 \\ l_2 & l_4 & d_3 & u_6 \\ l_3 & l_5 & l_6 & d_4 \end{array}$$

- If we want the relation to be reflexive, we must make all the $d_i = 1$.
- If we want the relation to be symmetric, then we must have $u_i = l_i$ for all i.
- If we want the relation to be antisymmetric, then we can never have $u_i = l_i = 1$ for any i (but $u_i = l_i = 0$ is allowed).

You should make sure you understand the reasons for these three statements. You may have noticed that transitivity was not mentioned. It cannot be described in such simple terms.

The 4×4 matrix above can easily be extended to a general $n \times n$ matrix. As we continue to describe certain properties of the 4×4 case, you should think about how these descriptions extend to the $n \times n$ case. As far as the situations just described, we do exactly the same thing in the $n \times n$ case.

From what we have just said, there are 2^{12} reflexive relations on a four element set. Why? We must set all of the $d_i = 1$ and then we can choose freely the u_i and l_i to be 0 or 1. There are 12 total u_i and l_i, giving 2^{12} choices. In general, there are $n^2 - n$ entries in an $n \times n$ matrix which are not d_i's. Thus there are $2^{n^2-n} = 2^{n(n-1)}$ reflexive relations on a set with n elements.

Let's try one more example. How many relations are both reflexive and antisymmetric? All of the $d_i = 1$. For each pair (l_i, u_i), we have three choices: $(l_i, u_i) = (0, 0)$, $(l_i, u_i) = (1, 0)$, $(l_i, u_i) = (0, 1)$. In the 4 element case there are 3^6 such relations. What is the formula for general n? We have seen that there are $n^2 - n$ elements l_i and u_i and so there are $(n^2 - n)/2$ pairs (l_i, u_i). Thus there are $3^{(n^2-n)/2}$ relations on an n-set which are both reflexive and antisymmetric. $\square$

Example 15 (Partitions of a set) The collection $\Pi(S)$ of all partitions of a set S can be made into a poset. Let $S = \underline{15}$. Consider the following partition of S:

$$\alpha = \Big\{\{1\},\ \{2\},\ \{9\},\ \{3,5\},\ \{4,7\},\ \{6,8,10,15\},\ \{11,12,13,14\}\Big\}.$$

We can *refine* the partition α by taking any block or blocks with at least two elements and splitting each of them into two or more blocks. For example, we could choose the block $\{6, 8, 10, 15\}$ and split it into two blocks: $\{6, 15\}, \{8, 10\}$. We could also choose the block $\{11, 12, 13, 14\}$ and split it into (for example) three blocks: $\{13\}, \{14\}, \{11, 12\}$. The resulting partition β is called a *refinement* of α (we write $\beta \preceq \alpha$):

$$\beta = \Big\{\{1\},\ \{2\},\ \{9\},\ \{3,5\},\ \{4,7\},\ \{6,15\},\ \{8,10\},\ \{13\},\ \{14\},\ \{11,12\}\Big\}.$$

The set of all partitions of S, $\Pi(S)$, together with the refinement relation is a poset — the lattice of partitions of S. By definition, $\alpha \preceq \alpha$ for any $\alpha \in \Pi(S)$. We leave it to you to check antisymmetry and transitivity. $\square$

New Posets from Old Ones

We now examine restrictions, direct products and lexicographic order, which are three ways of forming new posets from old ones.

Definition 5 (Restriction of a poset) *Let $(S, \preceq)$ be a poset and let X be a subset of S. For $u, v \in X$, define the order relation $\preceq_X$ by $u \preceq_X v$ if and only if $u \preceq v$ in $(S, \preceq)$. We call $(X, \preceq_X)$ a subposet of S or the restriction of S to X. Instead of introducing a new symbol $\preceq_X$ for the order relation, one usually uses $\preceq$, writing $(X, \preceq)$.*

Example 16 (Restrictions of posets) Consider the divides poset on $S = \underline{20}$. Let $X = \{2, 4, \ldots, 20\}$ — the even numbers in S. The set X with the same divides relation is a subposet of S or a restriction of S to X. Alternatively, let Y be the divisors of 20, namely $\{1, 2, 4, 5, 10, 20\}$ with the same divides relation. This is also a subposet.

As another example, consider the subset lattice $\mathcal{P}(S)$, $S = \{a, b, c\}$. Remove from this poset the empty set and the set S. This gives a new poset $\mathcal{P}'(S)$ with six elements (six subsets of S) ordered by set inclusion. The poset $\mathcal{P}'(S)$ is a subposet of $\mathcal{P}(S)$:

$$\mathcal{P}'(S) = \Big\{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\Big\}.$$

Since this is a restriction of $\mathcal{P}(S)$, it still has "subset of" as the order relation. □

Example 17 (Direct products of posets) Suppose we have two posets P and Q. Let $\preceq_P$ be the relation on P and $\preceq_Q$ be the relation on Q. The *direct product* of the posets $(P, \preceq_P)$ and $(Q, \preceq_Q)$ is the poset $(P \times Q, \preceq)$ where $(p_1, q_1) \preceq (p_2, q_2)$ if $p_1 \preceq_P p_2$ and $q_1 \preceq_Q q_2$. Sometimes this product order is called "coordinate order." Just as we can define Cartesian product of several sets, we can define the direct product of several posets.

A simple application of this idea is to take $P = Q = \{0, 1\}$. Then,

$$P \times Q = \{(0, 0),\ (0, 1),\ (1, 0),\ (1, 1)\}$$

has just four elements. Suppose $S = \{a, b\}$ is a two-element set. We can think of the elements of $P \times Q$ as the one-line forms for the functions $f : S \to \{0, 1\}$. With each of these functions is associated the subset $f^{-1}(1)$ of S. Using one-line notation for functions with S in the order (a, b), we see that

- $(0, 0)$ corresponds to the empty set,
- $(0, 1)$ corresponds to the set $\{b\}$,
- $(1, 0)$ corresponds to the set $\{a\}$, and
- $(1, 1)$ corresponds to the set $\{a, b\}$.

In this way the four functions $f : S \to \{0, 1\}$ become alternative descriptions of the four subsets of the subset lattice $\mathcal{P}(S)$. Thought of in this way, they are called the *characteristic functions* of the subsets of S. The poset of characteristic functions with coordinate order is just another way to describe the subset lattice with set inclusion.

Instead of a two element subset, we could consider an n-set S. In this case, we form the direct product of n copies of $P = (\{0, 1\}, \leq)$. The characteristic function is a bijection from $\mathcal{P}(S)$ to n-long vectors of zeroes and ones. We leave it to you to fill in the details. □

Is the direct product of posets again a poset? Yes. In fact we now define the direct product of binary relations and prove that properties are "inherited."

Definition 6 (Direct product of binary relations) *Let $S_1, S_2, \ldots, S_n$ be sets and let R_i be a binary relation on S_i for $i = 1, \ldots, n$. The direct product is the Cartesian product $S = S_1 \times S_2 \times \cdots \times S_n$ with the binary relation R defined by*

$$(a_1, a_2, \ldots, a_n)\ R\ (b_1, b_2, \ldots, b_n) \quad \text{if and only if } a_i\ R_i\ b_i \text{ for } i = 1, 2, \ldots, n.$$

You should verify that this definition gives the definition for the direct product of posets when the (S_i, R_i) are all posets.

The following theorem implies that the direct product of posets is again a poset. There are four statements in one — choose any property in the first { } and then choose the same property in the second { }

Theorem 4 (Properties of direct products) *If each of the binary relations R_i on the set S_i is* $\left\{\begin{matrix}\textit{reflexive}\\ \textit{symmetric}\\ \textit{antisymmetric}\\ \textit{transitive}\end{matrix}\right\}$*, then the direct product is also* $\left\{\begin{matrix}\textit{reflexive}\\ \textit{symmetric}\\ \textit{antisymmetric}\\ \textit{transitive}\end{matrix}\right\}$.

Proof: We prove transitivity and leave the rest to you. Suppose that

$$(a_1, a_2, \ldots, a_n)\ R\ (b_1, b_2, \ldots, b_n) \quad \text{and} \quad (b_1, b_2, \ldots, b_n)\ R\ (c_1, c_2, \ldots, c_n)$$

From the definition of R, $a_i\ R_i\ b_i$ and $b_i\ R_i\ c_i$ for $i = 1, \ldots, n$. Since R_i is transitive, $a_i\ R_i\ c_i$ for $i = 1, \ldots, n$. By the definition of R, $(a_1, a_2, \ldots, a_n)\ R\ (c_1, c_2, \ldots, c_n)$. This proves transitivity of R. □

Definition 7 (Isomorphism of posets) *Let $(S, \preceq_S)$ and $(T, \preceq_T)$ be posets. We say the posets are isomorphic if we have a bijection $f : S \to T$ such that $x \preceq_S y$ if and only if $f(x) \preceq_T f(y)$. We then call f an isomorhism between the posets.*

In Example 17, we used the characteristic function to construct an isomorphism between the subset lattice and the direct product $P \times \cdots \times P$ where $P = (\{0, 1\}, \leq)$. We now look at another example.

Example 18 (The divisibility relation again) Let $T = \{1, 2, 3, 4, 6, 12\}$, the set of divisors of 12, and let the order relation be "divides." Consider the two chains (linear orders) $C_2 = \{1, 2, 4\}$ and $C_3 = \{1, 3\}$ where the ordering can be thought of as either divisibility or ordinary $\leq$ since it gives the same ordering.

The posets T and $S = C_2 \times C_3$ describe the same situation. To see this, let $(a, b) \in C_2 \times C_3$ correspond to $ab \in S$. In this case $f((a, b)) = ab$ and the posets are isomorphic.

The previous idea can be applied to the set of divisors of n for any $n > 0$. The number of chains will equal the number of different primes dividing n.

We looked at the divisors of 12. What about the set $\underline{12}$ of positive integers less than or equal to 12 ordered by divisibility? This is not isomorphic to a direct product of chains. However, it is isomorphic to a restriction of a direct product of chains. Here is one way to do this. Let

$$V = (\{1, 2, 4, 8\}, |\,) \times (\{1, 3, 9\}, |\,) \times (\{1, 5\}, |\,) \times (\{1, 7\}, |\,) \times (\{1, 11\}, |\,).$$

The map $f((a, b, c, d, e)) = abcde$ shows that V is isomorphic to W the poset of divisors of $8 \times 9 \times 5 \times 7 \times 11$ with the divides relation. Since $\underline{12} \subset W$, $\underline{12}$ is a subposet W. □

The next concept, when applied to products of linearly ordered sets, is one of the most useful elementary ideas to be found in computer science. It's found in every subdiscipline of computer science and in almost every program of any length or complexity.

Definition 8 (Strings and lexicographic order) *Let $(S, \preceq)$ be a poset. We use S^* to denote the set of all "strings" over S; that is S^* contains*

- *for each $k > 0$, the set of k-long strings $(x_1, \ldots, x_k)$ of elements in S, which is denoted $S^k = \times^k S$;*
- *the empty string ϵ.*

We now define a relation $\preceq_L$ on S^*. This relation is called lexicographic order or, more briefly, lex order. Let $(a_1, a_2, \ldots, a_m)$ and $(b_1, b_2, \ldots, b_n)$ be two elements of S^* with $m, n > 0$. We say that

$$(a_1, a_2, \ldots, a_m) \preceq_L (b_1, b_2, \ldots, b_n)$$

if either of the following two conditions hold:

(1) $m \leq n$ and $a_i = b_i$ for $i = 1, \ldots, m$.

(2) For some $k < \min(m, n)$, $a_i = b_i$, $i = 1, \ldots, k$, $a_{k+1} \neq b_{k+1}$, and $a_{k+1} \preceq b_{k+1}$.

In addition we have a third condition:

(3) For the empty string ϵ, we have $\epsilon \preceq_L x$ for every string $x \in S^*$.

Notice that we said "define a relation $\preceq_L$ on S^*" rather than "define a partial order $\preceq_L$ on S^*." Why is that? By its definition, $\preceq_L$ is obviously a relation. The fact that it is a partial order requires proof. We'll give a proof after the next example.

The term “lexicographic” comes from the listing of words in a dictionary. One writes $S_0 = \{\epsilon\}$ and so $S^* = \cup_{k=0}^{\infty} S^k$. Other notations for the k-tuple $(x_1, x_2, \ldots, x_k)$ are $x_1, x_2, \ldots, x_k$ (leave off the parentheses) or $x_1 x_2 \ldots x_k$ (leave off the parentheses and the commas). Each of these latter notations leaves off information and can be confusing (for example, $112131212131414 = (11, 21, 312, 121, 31414)$, or does it?). Mostly, we will stick to the full notation (vector notation). Such k-tuples are sometimes referred to as *strings of length k over S*. The term “words” over S is also used in this context as meaning the same thing as “strings.”

Example 19 (Lexicographic order) At this stage, all we know is that $\preceq_L$ is a relation on strings. Consider the lattice of subsets of $\{1,2,3\}$; that is, $(\mathcal{P}(\{1,2,3\}), \subseteq)$. By Condition (1),

$$(\{1\},\{2\},\{1,3\}) \preceq_L (\{1\},\{2\},\{1,3\},\{1,2\}).$$

By Condition (2),

$$(\{1\},\{2\},\{1\},\{1,2\}) \preceq_L (\{1\},\{2\},\{1,3\}).$$

On the other hand,

$$(\{1\},\{2\},\{1,2\}) \not\preceq_L (\{1\},\{2\},\{1,3\})$$

and

$$(\{1\},\{2\},\{1,3\}) \not\preceq_L (\{1\},\{2\},\{1,2\}).$$

The two strings $(\{1\},\{2\},\{1,2\})$ and $(\{1\},\{2\},\{1,3\})$ are incomparable in lex order because $\{1,2\}$ and $\{1,3\}$ are incomparable in the subset lattice. If, on the other hand, S is linearly ordered (it is not in this example where S is the lattice of subsets) then S^* with lex order is linearly ordered. □

We now prove that the lexicographic relation in Definition 8 is an order relation on S^*.

Theorem 5 (The relation $\preceq_L$ is an order relation) *Let S be a set with order relation $\preceq$. Let $\preceq_L$ be the lexicographic relation on the strings S^*. Then, $\preceq_L$ is an order relation on S^*. If the poset S is linearly ordered, so is S^*.*

***Proof:** In the proof, we refer to the conditions (1), (2) and (3) of that discussion. We also omit commas and parentheses in the strings as there is no possibility of confusion.

First we show that the **reflexive** property is true. Let $w \in S^*$. If $w = \epsilon$ is the null string, then the reflexive property follows from Condition (3). If $w = a_1 a_2 \cdots a_m$ with $m \geq 1$ then the reflexive property follows immediately from Condition (1).

Next we show the **antisymmetric** property. Suppose that

$$a_1 a_2 \cdots a_m \preceq_L b_1 b_2 \cdots b_n \text{ and } b_1 b_2 \cdots b_n \preceq_L a_1 a_2 \cdots a_m.$$

This would be impossible if either of these relations were due to Condition (2) since $a_{k+1} \preceq b_{k+1}$ and $a_{k+1} \neq b_{k+1}$ implies $b_{k+1} \not\preceq a_{k+1}$. We have used here the fact that $\preceq$ is a partial order relation on S and hence is antisymmetric. Thus both relations are due to

Condition (1). Hence, $m \le n$ and $n \le m$, so $m = n$, and $a_i = b_i$ for $1 \le i \le m$. This proves the antisymmetric property.

Next we show the **transitive** property. Suppose that

$$a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n \quad\text{and}\quad b_1b_2\cdots b_n \preceq_L c_1c_2\cdots c_t.$$

If $m = 0$ (the empty string), transitivity is trivial by Condition (3). Thus we may suppose $m, n, t > 0$. We consider cases.

- If both $\preceq_L$ relations are due to Condition (1) then $a_1a_2\cdots a_m \preceq_L c_1c_2\cdots c_t$ from Condition (1) also.
- If $a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n$ is due to Condition (1) but $b_1b_2\cdots b_n \preceq_L c_1c_2\cdots c_t$ is due to Condition (2), let k be the smallest integer such that $b_{k+1} \ne c_{k+1}$. If $k < m$ then, using the fact that $\preceq$ is a partial order, $a_1a_2\cdots a_m \preceq_L c_1c_2\cdots c_t$ follows from Condition (2), otherwise (i.e., $k \ge m$) it follows from Condition (1).
- If $a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n$ is due to Condition (2) but $b_1b_2\cdots b_n \preceq_L c_1c_2\cdots c_t$ is due to Condition (1), the proof is similar to the preceding case.
- Finally, if $a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n$ and $b_1b_2\cdots b_n \preceq_L c_1c_2\cdots c_t$ are both due to Condition (2), then, using the fact that $\preceq$ is a partial order relation, $a_1a_2\cdots a_m \preceq_L c_1c_2\cdots c_t$ follows from Condition (2) also.

We have proved that $\preceq_L$ is an order relation.

It remains to show that if the poset S is **linearly ordered**, so is S^*. Suppose that we are given any two strings $a_1a_2\cdots a_m$ and $b_1b_2\cdots b_n$. Suppose, without loss of generality, that $m \le n$. Then either

(a) $a_1 = b_1, \ldots, a_m = b_m$ or

(b) there is a smallest $k < m$ such that $a_{k+1} \ne b_{k+1}$.

We consider cases.

- If (a) holds, then $a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n$ by Condition (1).
- If (b) holds, then either $a_1a_2\cdots a_m \preceq_L b_1b_2\cdots b_n$ or $b_1b_2\cdots b_n \preceq_L a_1a_2\cdots a_m$ by Condition (2). This follows since either $a_{k+1} \preceq b_{k+1}$ or $b_{k+1} \preceq a_{k+1}$, because $\preceq$ is a linear order.

We have shown that, given any two strings, either the first is less than or equal to the second in lex order or the reverse. Thus, S^* is linearly ordered by lex order. □

There is a variation on lex order on S^* which first orders the strings by length. Strings of the same length are then ordered lexicographically by restricting the above definition to subsets of S^* of the same length. This order relation is called *short lexicographic order* or *length-first lexicographic order*.

We now present an alternative proof of the theorem. You may find this proof harder to follow because it is somewhat more abstract. So why give it? It illustrates some techniques that are often used by people doing mathematics. Specifically, we will reduce an infinite problem (S^*) to a finite one (S_m below). We then embed this in another problem (T_m below) that lets us use previous results (Theorem 4).

Proof: Review Definitions 5 and 6 on restrictions and direct products of posets.

Let $S_m = \cup_{k=0}^m S^k$. We claim it suffices to prove the theorem for the restriction of $(S^*, \preceq_L)$ to $(S_m, \preceq_L)$ for each $m = 1, 2, \ldots$. Why? The conditions for being a partial order and for being a linear order refer to at most three elements of the poset. Suppose, for example, we want to verify transitivity. Someone gives us $r \preceq_L s$ and $s \preceq_L t$. Let m be the longest length of r, s and t. Then $r, s, t \in S_m$. By the definition of a restriction $\preceq_L$ is the same in the restriction as it is in the full poset, so we can work in the restriction. In other words, if the theorem is true with S^* replaced by S_m, then it is true for S^*.

We begin by adding a "blank" to S. As we shall see, this lets us use a restriction of a direct product. Let $\sqcup$ be something that is not in S. You can think $\sqcup$ as a blank space. Let $T = S \cup \{\sqcup\}$ and define $\preceq$ on T to be the same as it is on S together with $\sqcup \preceq x$ for all $x \in S$. We'll still call it $\preceq$. Let $(T_m, \preceq_m)$ be the direct product of m copies of $(T, \preceq)$. This is a poset by Theorem 4.

Let U_m be the restriction of T_m to those strings in which $\sqcup$ is never followed by an element of S, but may be followed by more $\sqcup$'s. In other words, blanks appear only at the end of a string.

Define $f(x)$ for $x \in S_m$ to be x "padded out" at the end with enough blanks to give a string of length m. You should have no trouble verifying that f is a bijection from S_m to U_m. You should also check that $x \preceq_L y$ if and only if $f(x) \preceq_m f(y)$. This gives us an isomorphism between $(S_m, \preceq_L)$ and the poset $(T_m, \preceq_m)$. Thus $(S_m, \preceq_L)$ is a poset. $\square$

Example 20 (Lexicographic bucket sort) Let $S = \{1, 2, 3\}$ be ordered in the usual way (as integers). Consider all strings of length three, $\times^3 S = S^3$. Take some subset of S^3, say the set

$$A = \big\{(2,1,3),\ (3,2,3),\ (1,2,1),\ (2,3,2),\ (1,1,3),\ (3,1,1),\ (3,3,1),\ (2,2,2)\big\}.$$

We are interested in an algorithm for sorting the elements of A so that they are in lexicographic order. The topic of sorting is very important for computer science. The literature on sorting methods is vast.

One type of sorting algorithm involves comparisons only. Imagine a bin of bananas that are to be sorted by weight using only a beam balance that tells which of two bananas weighs the most. No numerical values are recorded. Start with one banana. Get another and compare it with the first, laying them on a table in order of weight, left to right. Each new banana is compared with the bananas already sorted until all bananas are sorted by weight. In this manner you can sort the bananas by weight without actually knowing the numerical value of the weight of any banana. Such a sorting algorithm is called a *comparison sort.*

As another approach to sorting bananas by weight, suppose we have a scale that returns the weight of a banana to the nearest one-tenth of an ounce. Suppose we know that the bananas in the bin weigh between 4.0 oz. and 6.0 oz. We put 21 buckets on the table, the buckets labeled with the number 4.0, 4.1, 4.2, . . . , 5.7, 5.8, 5.9, 6.0. Take the bananas from the bin one by one and weigh them, accurate to 0.1 oz. Put each banana in the bucket corresponding to its weight. This type of sorting is called a *bucket sort*.

These two basic types of sorting, as well as hybrid forms of these two types, occur in many computer related applications. Comparison sorts are used with tree-type data structures, linked lists, etc. Bucket sorts are implemented by arrays where the buckets correspond to index references into an array.

Returning to the set A, we can sort the elements of A lexicographically using a variation on the bucket sort. We use three buckets, labeled 1, 2, and 3. On the first "pass" through the set A we place elements of

$$A = \big\{(2,1,3),\ (3,2,3),\ (1,2,1),\ (2,3,2),\ (1,1,3),\ (3,1,1),\ (3,3,1),\ (2,2,2)\big\}$$

into buckets according to the rightmost entry in that element:

PASS 1
Bucket 1: (1, 2, 1), (3, 1, 1), (3, 3, 1)
Bucket 2: (2, 3, 2), (2, 2, 2)
Bucket 3: (2,1,3), (3, 2, 3), (1, 1, 3)

Although A is a set and, technically, has no order, we now have imposed a linear order on A. This linear order is gotten, from the placement in the buckets, by reading the elements in the buckets from left to right, first from Bucket 1, then from Bucket 2 and finally from Bucket 3. We obtain the Pass 1 *concatenated order*:

$$(1,2,1),\ (3,1,1),\ (3,3,1),\ (2,3,2),\ (2,2,2),\ (2,1,3),\ (3,2,3),\ (1,1,3).$$

Notice that if you just read the third elements of each vector (string) you obtain, in order left to right, $1,1,1,2,2,3,3,3$. These strings of length one are in order.

We now do PASS 2. We go through the Pass 1 concatenated order, left to right, putting strings into buckets based on the value of their second-from-the-right coordinate (middle coordinate). In carrying this out, it is essential that the the order of the strings in each bucket is the correct order relative to PASS 1 concatenated order. The term "bucket" is not suggestive of order. Perhaps "sublist" would be better here, but we follow conventional terminology. Here is the composition of the buckets after PASS 2:

PASS 2
Bucket 1: (3, 1, 1), (2, 1, 3), (1, 1, 3)
Bucket 2: (1, 2, 1), (2, 2, 2), (3, 2, 3)
Bucket 3: (3, 3, 1), (2, 3, 2)

PASS 2 concatenated order is

$$(3,1,1),\ (2,1,3),\ (1,1,3),\ (1,2,1),\ (2,2,2),\ (3,2,3),\ (3,3,1),\ (2,3,2).$$

Note now that the list of all last two elements is in lexicographic order:

$$(1,1),\ (1,3),\ (1,3),\ (2,1),\ (2,2),\ (2,3),\ (3,1),\ (3,2).$$

Finally, we do PASS 3, putting strings into buckets according to first elements, retaining within each bucket the PASS 2 concatenated order.

PASS 3
Bucket 1: (1, 1, 3), (1, 2, 1)
Bucket 2: (2, 1, 3), (2, 2, 2), (2, 3, 2)
Bucket 3: (3, 1, 1), (3, 2, 3), (3, 3, 1)

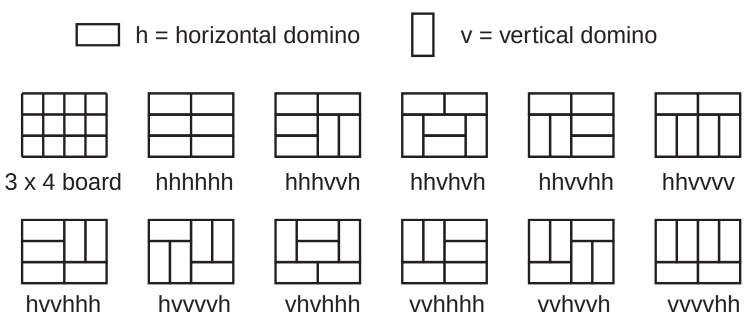
PASS 3 concatenated order is the lexicographic order on A.

$$(1,1,3),\ (1,2,1),\ (2,1,3),\ (2,2,2),\ (2,3,2),\ (3,1,1),\ (3,2,3),\ (3,3,1).$$

This process extends to a general algorithm called the *lexicographic bucket sort*. The correctness of the algorithm can be proved by induction on the number of passes (the length of the strings).

The lexicographic bucket sort was used in the early days of computers to sort "punched cards" on which data was stored. □

Example 21 (Lexicographic order and domino coverings) Lexicographic order is used in many computer applications where geometric objects are being manipulated as input data. We give an example from a class of problems called "tiling problems." Below we show a 3×4 "chess board" or grid. The grid is to be tiled or "covered" with horizontal and vertical dominoes (little 1×2 size rectangles):



If the squares of the board are numbered systematically, left to right, top to bottom, from 1 to 12, we can describe any placement of dominoes by a sequence of 6 h's and v's: Each of the domino placements in the above picture has such a description just below it. Take as an example, hhvhvh (the third domino covering in the picture). We begin with no dominoes on the board. None of the squares, numbered 1 to 12 are covered. The list of "unoccupied squares"is as follows:

1	2	3	4
5	6	7	8
9	10	11	12

Thus, the smallest unoccupied square is 1. The first symbol in hhvhvh is the h. That means that we take a horizontal domino and cover the square 1 with it. That forces us to cover square 2 also. The list of unoccupied squares is as follows:

		3	4
5	6	7	8
9	10	11	12

Now the smallest unoccupied square is 3. The second symbol in hhvhhv is also an h. Cover square 3 with a horizontal domino, forcing us to cover square 4 also. The list of unoccupied squares is as follows:

$$\begin{matrix} 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{matrix}$$

At this point, the first row of the board is covered with two horizontal dominoes (check the picture). Now the smallest unoccupied square is 5 (the first square in the second row). The third symbol in hhvhvh is v. Thus we cover square 5 with a vertical domino, forcing us to cover square 9 also. The list of unoccupied squares is as follows:

$$\begin{matrix} 6 & 7 & 8 \\ 10 & 11 & 12 \end{matrix}$$

We leave it to you to continue this process to the end and obtain the domino covering shown in the picture.

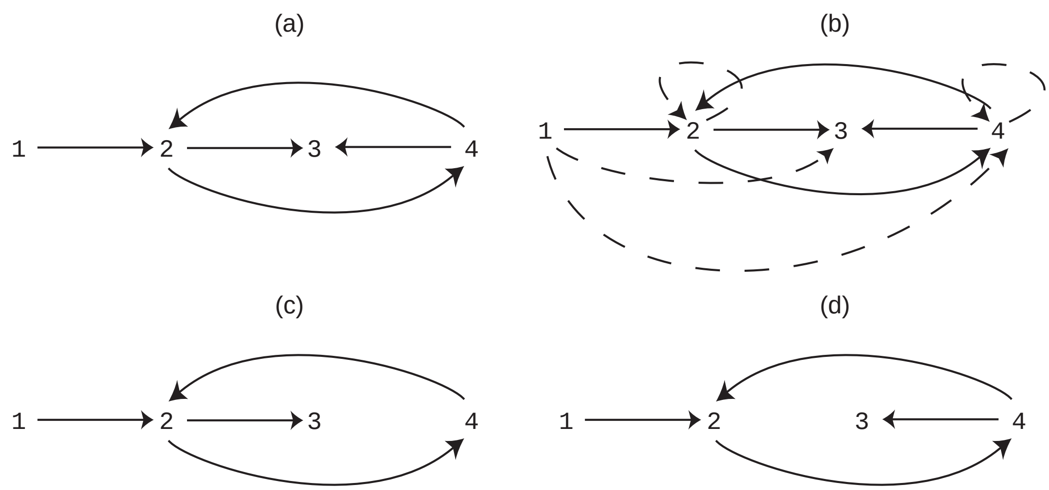
Here is the general description of the process. Place dominoes sequentially as follows. If the first unused element in the sequence is h, place a horizontal domino on the first (smallest numbered) unoccupied square and the square to its right. If the first unused element in the sequence is v, place a vertical domino on the first unoccupied square and the square just below it. Not all sequences correspond to legal placements of dominoes (try hhhhhv). For a 2×2 board, the only legal sequences are hh and vv. For a 2×3 board, the legal sequences are hvh, vhh and vvv. For a 3×4 board, there are eleven legal sequences as shown in the above picture.

Having developed this correspondence between tiling of a rectangular board and strings of letters from the set $S = \{h, v\}$, we can now list the strings that represent coverings of the board in lexicographic order. This order is useful for generating, storing, retrieving, and comparing domino coverings. □

More Poset Concepts

In the next example we introduce a useful pictorial or geometric way of visualizing a relation. The only problem with this method of thinking about relations is that the picture can become much too complicated. These pictures are combined with the idea of "transitive closure" to deal with this growth of complexity.

Example 22 (Transitive closure and directed graph diagrams) Consider the relation $R = \{(1,2), (2,3), (2,4), (4,3), (4,2)\}$ on the set $S = \{1,2,3,4\}$. Figure (a) below is another way of representing R using a *directed graph diagram*. The elements of S are written down in some manner, in this case one after the other in a straight line (any way will do). An arrow is drawn between i and j if and only if $(i,j) \in R$. Note that R is not transitive. For example, $(1,2) \in R$ and $(2,3) \in R$, but $(1,3) \notin R$. Suppose we add the missing pair $(1,3)$ to R as indicated in Figure (b) below. This gives a new relation $\{(1,2), (2,3), (2,4), (4,3), (4,2), (1,3)\}$. This new relation is still not transitive. For example, $(2,4) \in R$ and $(4,2) \in R$, but $(2,2) \notin R$. So, add $(2,2)$ to R. Keep repeating this process until no violations of transitivity can be found. The directed graph diagram of this final transitive relation is shown in Figure (b). This relation is the smallest transitive relation that contains R. It is called the *transitive closure* of R.



Note that there are even smaller (fewer elements) relations than R whose transitive closure is the same as R. The directed graph diagrams of two such relations are shown in Figures (c) and (d). By the way, we use the terminology "directed graph diagram," so you might naturally wonder what is a directed graph (the thing that is being "diagrammed" here)? The answer is that directed graphs and binary relations are, mathematically, the same thing. The terminology "directed graph diagram" is standard in this context, rather than the more natural "relation diagram." $\square$

***Example 23 (Transitive closure and matrices)** There is another way to compute transitive closure. We start the same relation, $R = \{(1,2), (2,3), (2,4), (4,3), (4,2)\}$ on the set $S = \{1,2,3,4\}$, used in the previous example. Here is the "incidence matrix" of this relation:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

The interpretation is that if $A(i,j) = 1$ then $(i,j) \in R$, else $(i,j) \notin R$. If we compute the square of A, then, by definition of matrix multiplication, $A^2(i,j) = \sum_{k=0}^{4} A(i,k)A(k,j)$. Note that $A^2(i,j) \neq 0$ if and only if there is a pair $(i,t) \in R$ and $(t,j) \in R$. In other words, for this pair, $A(i,t)A(t,j) = 1$ so $A^2(i,j) > 0$. For our purposes, we don't care how big

$A^2(i,j)$ is only whether or not it is zero. So, we replace all nonzero entries in $A^2(i,j)$ by 1. This is called the "Boolean product" of A with A. We just use the same notation for this Boolean product as for the square. Here is this Boolean product:

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

You can think of the Boolean product as follows: Multiply matrices in the usual way be replace "plus" with $\vee$ and "times" with $\wedge$. Thus

$$A^2(i,j) = \big(A(i,1) \wedge A(1,j)\big) \vee \big(A(i,2) \wedge A(1,2)\big) \vee \cdots \vee \big(A(i,n) \wedge A(1,n)\big).$$

If we now form the "Boolean sum" $A + A^2$ (Again, replace "plus" with $\vee$.) we get the following matrix:

$$A + A^2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

This matrix has a 1 in position (i,j) if and only if either $(i,j) \in R$ or $(i,t) \in R$ and $(t,j) \in R$ for at least one t (perhaps both of these conditions hold). In terms of the directed graph diagram, this matrix has a 1 in position (i,j) if and only if there is a directed arrow joining i to j or a sequence of two directed arrows that you can follow to go from i to j. Such a sequence is called a directed path of length two from i to j.

We could continue this process to compute the Boolean matrix $A + A^2 + A^3$, but, if we do, we see that this latter matrix is the same as $A + A^2$. A little thought should tell you that this means $A + A^2$ is the incidence matrix of the transitive closure of R.

This idea can be applied to any binary relation. For large relations a computer helps. You start with the incidence matrix A and keep forming Boolean partial sums $S_k = \sum_{i=1}^k A^i$ until, for some $k = t$, $S_t = S_{t+1}$. Then we'll have $S_{t+i} = S_t$ for all $i \geq 0$. At this point S_t is the incidence matrix of the transitive closure. Note that $S_1 = A$, $S_2 = A\,S_1 + A$, $S_3 = A\,S_2 + A$, and, in general $S_{k+1} = A\,S_k + A$. This is a convenient way to carry out these computations. It also makes it easy to prove the earlier claim that $S_{t+i} = S_t$: Use induction in i and note that $S_{t+i} = A\,S_{t+i-1} + A$, which equals $A\,S_t + A$ by the induction hypothesis.

The story doesn't end here. We can find a similar algorithm that is much faster for large problems. We claim that $S_t\,S_t + A = S_{2t}$. Why is this? If you consider ordinary multiplication, you should see that $S_t\,S_t$ consists of all the powers $A^2, A^3, \ldots, A^{2t}$ added together, some of them many times. Notice that when you do Boolean addition for any matrix B, you have $B + \cdots + B = B$. Thus $S_t\,S_t = A^2 + \cdots + A^{2t}$ when we do Boolean addition and multiplication. Here's our new algorithm:

$$P_0 = A \quad \text{and} \quad P_{k+1} = P_k\,P_k + A \quad \text{for} \quad k \geq 0,$$

and we stop when $P_k = P_{k+1}$. You should be able to prove by induction that $P_k = S_{2^k}$.

Is this algorithm really faster? Yes. The simplest example of this is a chain:

$$S = \{0, 1, \ldots, n\} \quad \text{and} \quad R = \{(0,1), (1,2), \ldots, (n-1,n)\}.$$

We claim that $S_{n+1} = S_n$ and the n matrices $S_1, S_2, \ldots, S_n$ are all different.[4] Thus we must compute $S_2, \ldots, S_{n+1}$ to obtain the transitive closure if we use the first algorithm. For the second algorithm, we compute $P_1, \ldots, P_{m+1}$ where $m = \lceil \log_2 n \rceil$. Why this value of m? We will not have $P_k = P_{k+1}$ until $2^k \geq n$ and m is the smallest such k. For large values of n, $\log_2 n$ is *much* smaller than n.

Which algorithm should you use in the problems? It doesn't make much difference because the sets S we look at are small. $\square$

Example 24 (Covering relations and Hasse diagrams) Let S be a finite poset with relation $\preceq$. We define a new relation on S called the *covering relation*, denoted by $\prec_c$. For $x, y \in S$, we say $x \prec_c y$ if

(a) $x \neq y$ and $x \preceq y$, and

(b) $x \preceq z \preceq y$ implies that either $x = z$ or $y = z$.

In words, $x \prec_c y$ if x and y are different and there is no third element of S "between" x and y. In this case, we say that "y covers x" or "x is covered by y."

The condition $x \neq y$ means that the covering relation of a nonempty set S is never reflexive. In fact, $x \not\prec_c x$ for all $x \in S$. Thus, the covering relation fails badly the test of being reflexive. The covering relation is always antisymmetric. If there are three distinct elements $a \preceq b \preceq c$ in S then the covering relation is not transitive; otherwise, it is trivially transitive.

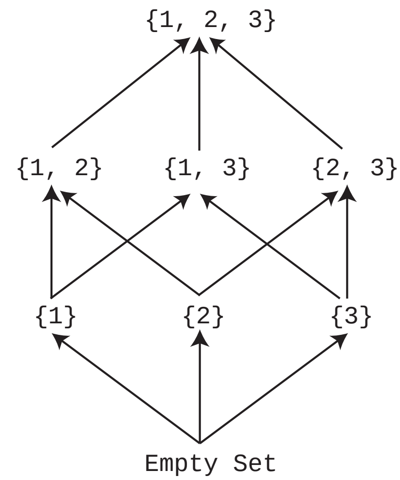
If you recall the discussion of Example 22 you can easily see that any order relation is *almost* the transitive closure of its covering relation. Missing, when we take the transitive closure of the covering relation, are all of the relations of the form $x \preceq x$. If you add those at the end, after taking the transitive closure of the covering relation, then you recover $\preceq$. Or, start with the covering relation, add in all pairs (x, x), $x \in S$, and then take the transitive closure.[5]

Take for example a set $A = \{1, 2, 3\}$ and the subset lattice $\mathcal{P}(A)$. Let $x = \{1\}$ and $y = \{1, 2, 3\}$. In this example, $x \subseteq y$ but y does not cover x, written $x \not\subset_c y$. To see why, note that there is a third element $z = \{1, 2\}$ between x and y: $\{1\} \subseteq \{1, 2\} \subseteq \{1, 2, 3\}$. In this example, x is covered by z and z is covered by y. Here is the directed graph diagram

[4] You are encouraged to experiment with small values of n to convince yourself that this is true.

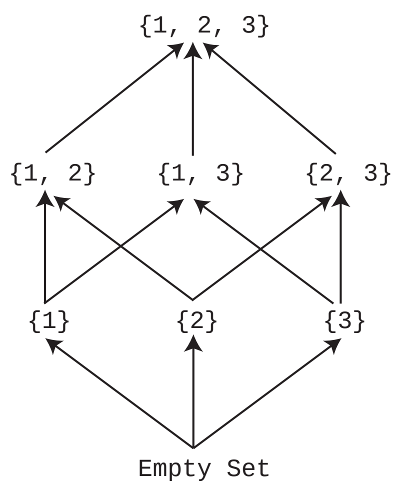
[5] If S is not finite, this process may not work. For example, the covering relation for the real numbers, $(\mathbb{R}, \leq)$, is empty!

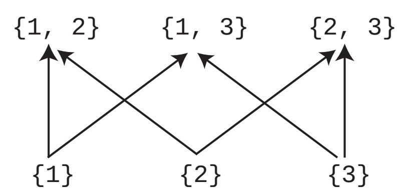
for the covering relation:



The directed graph diagram of the covering relation of a poset is called the *Hasse diagram* of the poset. The Hasse diagram is a very useful geometric way to picture posets. The transitive closure of the relation represented by the Hasse diagram, plus all pairs (x, x), x in the poset, is the order relation. □

Example 25 (Poset terminology) Here are the Hasse diagrams of two posets, the second a subposet of the first:



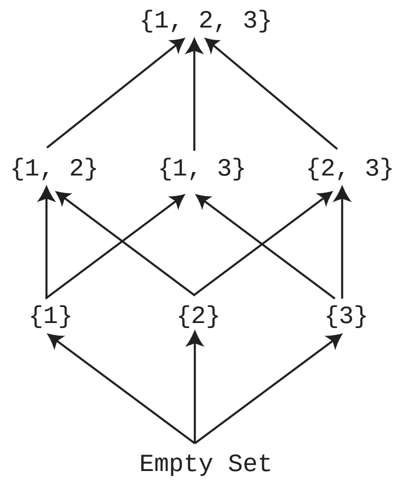


The subset $\big\{\{1\}, \{1,3\}, \{1,2,3\}\big\}$, of the first poset is called a *chain* in that poset. It is a chain because, as a subposet, it is linearly ordered: $\{1\} \subseteq \{1,3\} \subseteq \{1,2,3\}$. The *length* of this chain is two (one less than the number of elements in the chain). The longest chain in this poset has length three. There are six such "maximal" chains. You should try to find them all. In this first poset, the empty set $x = \emptyset$ is special in that for all y in the poset, $x \subseteq y$. Such an element is called the *least element* in the poset. Correspondingly, the element $t = \{1,2,3\}$, is the *greatest element* in the poset because $y \subseteq t$ for all y in the poset. There can be at most one greatest element and at most one least element in a poset.

Consider now the second poset. There is no least element and no greatest element in this poset. The element $x = \{2\}$ has the property that there is no y in the poset with $y \neq x$

and $y \subseteq x$. Such an element x is called a *minimal element* of a poset. A least element is a minimal element, but not necessarily the other way around. Similarly, $\{1, 2\}$ is a *maximal element* of this poset (but not a greatest element). Confused? Read it over again and look at the pictures. It is not that bad! □

Example 26 (Linear extensions — topological sorts) For this example, we shall return to the lattice of subsets of the set $\{1, 2, 3\}$. Recall its Hasse diagram:



Here is a special listing of the elements of the lattice of subsets in which every element occurs exactly once:

$$\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}.$$

What is special about this listing? If you scan the list from left to right, you will find that for each set in the list, all of its supersets are to the right of it. Or, said in another way, if you scan from right to left, no set is a subset of some other set located to its left. Such a listing is called a *linear extension* of the poset (by mathematicians) or a *topological sort* (by computer scientists). Here is another linear extension of the poset:

$$\emptyset, \{3\}, \{2\}, \{1\}, \{1,3\}, \{1,2\}, \{2,3\}, \{1,2,3\}.$$

This poset has 48 linear extensions. Can you list them all?

Here is a listing of the elements of the lattice of subsets that is not a linear extension:

$$\emptyset, \{1\}, \{2\}, \{1,2\}, \{1,3\}, \{3\}, \{2,3\}, \{1,2,3\}.$$

Scanning from left to right, the set $\{3\}$ does not have all of its supersets to the right of it (the superset $\{1, 3\}$ is to the left). Scanning from right to left, we again see that $\{3\}$ is a subset of $\{1, 3\}$ which is to the left.

In general, let S be a poset with n elements and with relation $\preceq$. A *linear extension* of S is a listing of the elements of S, $s_1, s_2, \ldots, s_n$, such that for any $1 \le i, j \le n$, if $s_i \preceq s_j$ then $i \le j$.

It is usually difficult to count the number of linear extensions of an arbitrary poset except by listing them. There are no easy formulas for many common posets. An easy case is an n-set with the empty relation (the "discrete" poset): There are $n!$ linear extensions. □

Exercises for Section 2

2.1. In each case a binary relation R on a set S is specified directly as a subset of $S \times S$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

(a) $R = \{(0,0), (0,1), (0,3), (1,0), (1,1), (2,3), (3,3)\}$ where $S = \{0,1,2,3\}$.

(b) $R = \{(1,3), (3,1), (0,3), (3,0), (3,3)\}$ where $S = \{0,1,2,3\}$.

(c) $R = \{(a,a), (a,b), (b,c), (a,c)\}$ where $S = \{a,b,c\}$.

(d) $R = \{(a,a), (b,b)\}$ where $S = \{a,b,c\}$.

(e) $R = \emptyset$ where $S = \{a\}$.

2.2. Define a binary relation on $\mathbb{R}$ (the reals) by $x\ R\ y$ if $\exists\, n \in \mathbb{Z}$ (the integers) such that $x^2 + y^2 = n^2$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.3. Define a binary relation on $\mathbb{Z}$ by $x\ R\ y$ if $x = y$ or if $x - y = 2k + 1$ for some integer k. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.4. Let $S = \mathbb{R}$, the real numbers. Define a binary relation on S by $x\ R\ y$ if $x^2 = y^2$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.5. Define a binary relation on $\mathbb{N}^+$ (the positive integers) by $x\ R\ y$ if $\gcd(x, y) > 1$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.6. Let $S = \mathcal{P}(\underline{4}) - \{\emptyset\}$, the power set of $\underline{4} = \{1,2,3,4\}$ with the empty set discarded. Define a binary relation on S by $X\ R\ Y$ if $X \cap Y \neq \emptyset$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.7. Let $S = \mathcal{P}(T)$, be the power set of $T = \{1,2,3,4\}$. Define a binary relation on S by $X\ R\ Y$ if either $X \subseteq Y$ or $Y \subseteq X$. Determine, for each property, whether the relation R is reflexive, symmetric, or transitive. Explain your answers.

2.8. Let S be a set with n elements. How many binary relations on S are reflexive? How many are not reflexive?

2.9. Let S be a set with n elements. How many binary relations on S are symmetric? How many are not symmetric?

Equivalence and Order

2.10. Let S be a set with n elements. How many binary relations on S are reflexive and symmetric?

2.11. Let S be a set with n elements. How many binary relations on S are antisymmetric? How many are reflexive and antisymmetric?

2.12. Let $R = \{(0,0),(0,3),(1,0),(1,2),(2,0),(3,2)\}$ be a binary relation on $\{0,1,2,3\}$. Find the transitive closure of R.

2.13. Let $R = \{(a,c),(b,c),(c,d)\}$ be a binary relation on $\{a,b,c,d\}$. Find the transitive closure of R by experimentation and by the matrix method.

2.14. Let S be the set of composite integers n, $4 \leq n \leq 20$. Order S with the divides relation. What is the covering relation? Draw the Hasse diagram. List the minimal and maximal elements. Specify a chain of longest length.

2.15. Let $S = \{1,2,3,4,5\}$. Let $\mathcal{P}^{(2)}(S)$ denote the subset of $\mathcal{P}(S)$ consisting of all subsets A such that if $i,j \in A$, then $i \neq j$ implies that $|i-j| \geq 2$. Order the elements of $\mathcal{P}^{(2)}(S)$ by set inclusion. What is the cardinality of the covering relation of $\mathcal{P}^{(2)}(S)$? How many chains are there of length three? What are the maximal elements? the minimal elements? Is there a greatest element? a least element?

2.16. Give an example of a poset with no maximal element.

2.17. Let $S_2 = \{0,1\} \times \{0,1\} = \times^2\{0,1\}$. Use coordinate order: $(x_1,x_2) \leq (y_1,y_2)$ if $x_1 \leq y_1$ and $x_2 \leq y_2$. What is the covering relation? Compare this covering relation with $\mathcal{P}(X)$, $|X| = 2$, and set inclusion as the order relation. How do these ideas extend to $\times^3\{0,1\}$. To $\times^n\{0,1\}$?

2.18. Let S be the set of composite integers n, $4 \leq n \leq 20$. Order S with the divides relation. Let S^* denote the set of all finite strings (words) over S ordered lexicographically based on the poset S. Answer the following by stating whether or not the pair of strings (read left to right) is in order, in reverse order, or incomparable lexicographically.

(a) 4 6 18 and 4 6 9

(b) 4 6 8 and 4 6 8 9

(c) 4 16 8 and 4 6 10 9

2.19. Let S be an n-element set where $n \geq 3$ and let $x, y \in S$ where $x \neq y$. Suppose that S is made into a poset in such a way that all pairs of elements are comparable except x and y.

(a) What is the covering relation for S and does the Hasse diagram of S look like? (Describe *all* possible answers.)

(b) Let the poset T be $S \times S$ with the lex order. How many pairs $\{(a_1, a_2),\ (b_1, b_2)\}$ of incomparable elements does T have?

2.20. List in lexicographic order all ways of placing six dominoes on a 2×6 board.

2.21. Sort the following list into lexicographic order using a three-pass bucket sort: 321, 441, 143, 312, 422, 221, 214, 311, 234, 111. (Each element in the list is a sequence of three digits — not a 3-digit number. Thus 321 is the list (3,2,1).) Show the composition of the buckets after each pass.

2.22. Let S be the set of composite integers n, $4 \le n \le 20$. Order S with the divides relation. Let $x_1, x_2, \ldots, x_{11}$ be a topological sort of this poset. A pair (i, j), where $i < j$ and the integer x_i is smaller than the integer x_j will be called an "in-order pair." Find a topological sort where the number of in order pairs is less than or equal to 26.
Hint: First draw the Hasse diagram.

2.23. Let $S = \{a, b, c\}$, a set with three elements. Let $\mathcal{P}(S)$ be the set of all subsets of S ordered by set inclusion. Find 48 different topological sorts of $\mathcal{P}(S)$. You need not list them all if you can describe them in a convincing way.

Multiple Choice Questions for Review

In each case there is one correct answer (given at the end of the problem set). Try to work the problem first without looking at the answer. Understand both why the correct answer is correct and why the other answers are wrong.

1. Let $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. What is the smallest integer K such that any subset of S of size K contains two disjoint subsets of size two, $\{x_1, x_2\}$ and $\{y_1, y_2\}$, such that $x_1 + x_2 = y_1 + y_2 = 9$?

(a) 8 (b) 9 (c) 7 (d) 6 (e) 5

2. There are K people in a room, each person picks a day of the year to get a free dinner at a fancy restaurant. K is such that there must be at least one group of six people who select the same day. What is the smallest such K if the year is a leap year (366 days)?

(a) 1829 (b) 1831 (c) 1830 (d) 1832 (e) 1833

3. A mineral collection contains twelve samples of Calomel, seven samples of Magnesite, and N samples of Siderite. Suppose that the smallest K such that choosing K samples from the collection guarantees that you have six samples of the same type of mineral is $K = 15$. What is N?

(a) 6 (b) 2 (c) 3 (d) 5 (e) 4

4. What is the smallest $N > 0$ such that any set of N nonnegative integers must have two distinct integers whose sum or difference is divisible by 1000?

(a) 502 (b) 520 (c) 5002 (d) 5020 (e) 52002

5. Let $S = \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21\}$. What is the smallest integer $N > 0$ such that for any set of N integers, chosen from S, there must be two distinct integers that divide each other?

(a) 10 (b) 7 (c) 9 (d) 8 (e) 11

6. The binary relation $R = \{(0, 0), (1, 1)\}$ on $A = \{0, 1, 2, 3, \}$ is

(a) Reflexive, Not Symmetric, Transitive

(b) Not Reflexive, Symmetric, Transitive

(c) Reflexive, Symmetric, Not Transitive

(d) Reflexive, Not Symmetric, Not Transitive

(e) Not Reflexive, Not Symmetric, Not Transitive

7. Define a binary relation $R = \{(0, 1), (1, 2), (2, 3), (3, 2), (2, 0)\}$ on $A = \{0, 1, 2, 3\}$. The directed graph (including loops) of the transitive closure of this relation has

(a) 16 arrows

(b) 12 arrows

(c) 8 arrows

(d) 6 arrows

(e) 4 arrows

8. Let $\mathbb{N}^+$ denote the nonzero natural numbers. Define a binary relation R on $\mathbb{N}^+ \times \mathbb{N}^+$ by $(m,n)R(s,t)$ if $\gcd(m,n) = \gcd(s,t)$. The binary relation R is

(a) Reflexive, Not Symmetric, Transitive

(b) Reflexive, Symmetric, Transitive

(c) Reflexive, Symmetric, Not Transitive

(d) Reflexive, Not Symmetric, Not Transitive

(e) Not Reflexive, Not Symmetric, Not Transitive

9. Let $\mathbb{N}_2^+$ denote the natural numbers greater than or equal to 2. Let mRn if $\gcd(m,n) > 1$. The binary relation R on $\mathbb{N}_2$ is

(a) Reflexive, Symmetric, Not Transitive

(b) Reflexive, Not Symmetric, Transitive

(c) Reflexive, Symmetric, Transitive

(d) Reflexive, Not Symmetric, Not Transitive

(e) Not Reflexive, Symmetric, Not Transitive

10. Define a binary relation R on a set A to be *antireflexive* if xRx doesn't hold for any $x \in A$. The number of symmetric, antireflexive binary relations on a set of ten elements is

(a) 2^{10} (b) 2^{50} (c) 2^{45} (d) 2^{90} (e) 2^{55}

11. Let R and S be binary relations on a set A. Suppose that R is reflexive, symmetric, and transitive and that S is symmetric, and transitive but is **not** reflexive. Which statement is always true for any such R and S?

(a) $R \cup S$ is symmetric but not reflexive and not transitive.

(b) $R \cup S$ is symmetric but not reflexive.

(c) $R \cup S$ is transitive and symmetric but not reflexive

(d) $R \cup S$ is reflexive and symmetric.

(e) $R \cup S$ is symmetric but not transitive.

12. Define an equivalence relation R on the positive integers $A = \{2, 3, 4, \ldots, 20\}$ by $m\ R\ n$ if the largest prime divisor of m is the same as the largest prime divisor of n. The number of equivalence classes of R is

(a) 8 (b) 10 (c) 9 (d) 11 (e) 7

Equivalence and Order

13. Let $R = \{(a,a),(a,b),(b,b),(a,c),(c,c)\}$ be a partial order relation on $\Sigma = \{a,b,c\}$. Let $\preceq$ be the corresponding lexicographic order on Σ^*. Which of the following is true?

(a) $bc \preceq ba$

(b) $abbaaacc \preceq abbaab$

(c) $abbac \preceq abb$

(d) $abbac \preceq abbab$

(e) $abbac \preceq abbaac$

14. Consider the divides relation, $m \mid n$, on the set $A = \{2,3,4,5,6,7,8,9,10\}$. The cardinality of the covering relation for this partial order relation (i.e., the number of edges in the Hasse diagram) is

(a) 4 (b) 6 (c) 5 (d) 8 (e) 7

15. Consider the divides relation, $m \mid n$, on the set $A = \{2,3,4,5,6,7,8,9,10\}$. Which of the following permutations of A is **not** a topological sort of this partial order relation?

(a) 7,2,3,6,9,5,4,10,8

(b) 2,3,7,6,9,5,4,10,8

(c) 2,6,3,9,5,7,4,10,8

(d) 3,7,2,9,5,4,10,8,6

(e) 3,2,6,9,5,7,4,10,8

16. Let $A = \{2,3,4,5,6,7,8,9,10,11,12,13,14,15,16\}$ and consider the divides relation on A. Let C denote the length of the maximal chain, M the number of maximal elements, and m the number of minimal elements. Which is true?

(a) $C = 3$, $M = 8$, $m = 6$

(b) $C = 4$, $M = 8$, $m = 6$

(c) $C = 3$, $M = 6$, $m = 6$

(d) $C = 4$, $M = 6$, $m = 4$

(e) $C = 3$, $M = 6$, $m = 4$

Answers: **1** (c), **2** (b), **3** (e), **4** (a), **5** (d), **6** (b), **7** (a), **8** (b), **9** (a), **10** (c), **11** (d), **12** (a), **13** (b), **14** (e), **15** (c), **16** (a).