Unit SF

Sets and Functions

Section 1: Sets

The basic concepts of sets and functions are topics covered in high school math courses and are thus familiar to most university students. We take the intuitive point of view that sets are unordered collections of objects. We first recall some standard terminology and notation associated with sets. When we speak about sets, we usually have a "universal set" U in mind, to which the various sets of our discourse belong.

Definition 1 (Set notation) *A set is an unordered collection of distinct objects. We use the notation $x \in S$ to mean "x is an element of S" and $x \notin S$ to mean "x is not an element of S." Given two subsets (subcollections) of U, X and Y, we say "X is a subset of Y," written $X \subseteq Y$, if $x \in X$ implies that $x \in Y$. Alternatively, we may say that "Y is a superset of X." $X \subseteq Y$ and $Y \supseteq X$ mean the same thing. We say that two subsets X and Y of U are equal if $X \subseteq Y$ and $Y \subseteq X$. We use braces to designate sets when we wish to specify or describe them in terms of their elements: $A = \{a, b, c\}$, $B = \{2, 4, 6, \ldots\}$. A set with k elements is called a k-set or set with cardinality k. The cardinality of a set A is denoted by $|A|$.*

Since a set is an unordered collection of distinct objects, the following all describe the same 3-element set

$$\{a, b, c\} = \{b, a, c\} = \{c, b, a\} = \{a, b, b, c, b\}.$$

The first three are simply listing the elements in a different order. The last happens to mention some elements more than once. But, since a set consists of distinct objects, the elements of the set are still just a, b, c. Another way to think of this is:

> Two sets A and B are equal if and only if every element of A is an element of B and every element of B is an element of A.

Thus, with $A = \{a, b, c\}$ and $B = \{a, b, b, c, b\}$, we can see that everything in A is in B and everything in B is in A. You might think "When we write a set, the elements are in the order written, so why do you say a set is not ordered?" When we write something down we're stuck — we have to list them in some order. You can think of a set differently: Write each element on a separate slip of paper and put the slips in a paper bag. No matter how you shake the bag, it's still the same set.

If we are given that A is a set and no other information about A, then there is no ordering to the elements of A. Thus, we cannot speak of "the second element of the set A" unless we have specified an ordering of the elements of A. If we wish to regard A as ordered in some way, then we specify this fact explicitly: "The elements of A are ordered a, b, c," or "$A = (a, b, c)$." The latter notation replaces the braces with parentheses and designates that A is ordered, left to right, as indicated. We call this an *ordered set*. An ordered set is also called a *linear order*. Various other names are also used: *list*, *vector*, *string*, *word*

— all with **no repeated elements**.[1] Of course, you've seen repeated elements in vectors, for example the point in the plane at the coordinates (1,1). That's fine, it's just not an ordered set. If there are k elements in the ordered set, it is referred to as a *k-list, k-vector*, etc., or as a list, vector, etc., of *length k* — all with no repeated elements because they are ordered **sets**.

Sometimes we cannot list the elements of a set explicitly. What do we do if we want to describe the set of all real numbers greater than 1 without writing it out in words? We write

$$\{x \mid x \in \mathbb{R},\ x > 1\} \quad \text{or} \quad \{x \mid x > 1\} \quad \text{or} \quad \{x : x > 1\}.$$

These are read "the set of all x such that ..." In the first example we mentioned that x was a real number ($x \in \mathbb{R}$). In the other two we didn't because we assumed the reader knew from context that we were talking about real numbers.

For the most part, we shall be dealing with finite sets. Let U be a set and let A and B be subsets of U. The sets

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

and

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

are the *intersection* and *union* of A and B. The set $A \setminus B$ or $A - B$ is the *set difference* of A and B (i.e., the set $\{x \mid x \in A, x \notin B\}$). The set $U \setminus A$ (also A^c, A' or $\sim A$) is the *complement* of A (relative to U). Note that $A - B = \{x \mid x \in A, x \notin B\} = A \cap B^c$. The empty set, denoted by $\emptyset$, equals U^c. Also note that, for any set $A \subseteq U$, $A \cup A^c = U$ and $A \cap A^c = \emptyset$.

The set $A \oplus B = (A \setminus B) \cup (B \setminus A)$ is the *symmetric difference* of A and B. We use $A \times B = \{(x, y) \mid x \in A, y \in B\}$ to denote the *product* or *Cartesian product* of A and B. If we want to consider the product of k sets, $A_1, \ldots, A_k$, this is denoted by $\times_{i=1}^k A_i$. If we want to consider the product of a set A with itself k times, we write $\times^k A$.

Set Properties and Proofs

The algebraic rules for operating with sets are also familiar to most beginning university students. Here is such a list of the basic rules. In each case the standard name of the rule is given first, followed by the rule as applied first to $\cap$ and then to $\cup$.

Theorem 1 (Algebraic rules for sets) *The universal set U is not mentioned explicitly but is implicit when we use the notation $\sim X = U - X$ for the complement of X. An*

[1] Why is it okay to specify a set $S = \{a, b, c, a\}$ where the element a has been repeated, but it is not okay to have repeated elements in an ordering of S? When we say $S = \{a, b, c, a\}$, we know that S contains just the three elements a, b and c. If we were to talk about the ordered set (a, b, c, a) it would not make sense because it would say that the element a is in two places at once: the first position and the last position.

alternative notation is $X^c = \sim X$.

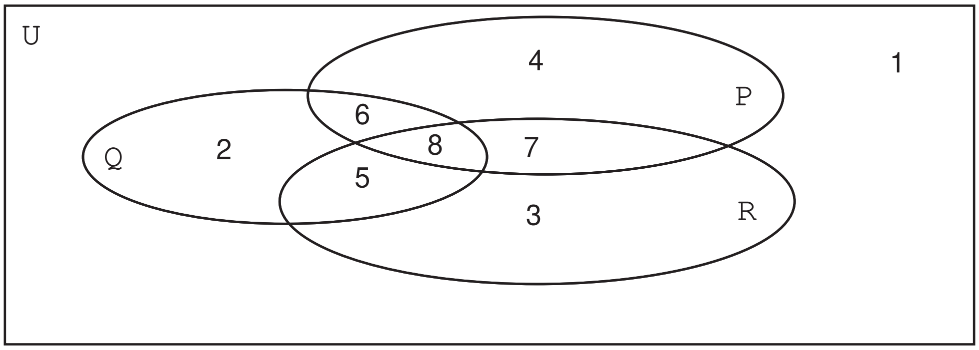
Associative:	$(P \cap Q) \cap R = P \cap (Q \cap R)$	$(P \cup Q) \cup R = P \cup (Q \cup R)$
Distributive:	$P \cap (Q \cup R) = (P \cap Q) \cup (P \cap R)$	$P \cup (Q \cap R) = (P \cup Q) \cap (P \cup R)$
Idempotent:	$P \cap P = P$	$P \cup P = P$
Double Negation:	$\sim\sim P = P$	
DeMorgan:	$\sim(P \cap Q) = \sim P \cup \sim Q$	$\sim(P \cup Q) = \sim P \cap \sim Q$
Absorption:	$P \cup (P \cap Q) = P$	$P \cap (P \cup Q) = P$
Commutative:	$P \cap Q = Q \cap P$	$P \cup Q = Q \cup P$

These rules are "algebraic" rules for working with $\cap$, $\cup$, and $\sim$. You should memorize them as you use them. They are used just like rules in ordinary algebra: whenever you see an expression on one side of the equal sign, you can replace it by the expression on the other side.

When we wrote "$P \cap Q \cap R$" you may have wondered if we meant "$(P \cap Q) \cap R$" or "$P \cap (Q \cap R)$." The associative law says it doesn't matter. That is why you will see the notation $P \cap Q \cap R$ or $P \cup Q \cup R$ without anyone getting excited about it. On the other hand $P \cap (Q \cup R)$ and $(P \cap Q) \cup R$ may not be equal, so we need parentheses here.

The best way to "prove" the rules or to understand their validity is through the geometric device of a *Venn diagram.*

Example 1 (Venn diagrams and proofs of set equations) Here is a Venn diagram for three sets, P, Q, and R, with universal set U:



The three oval regions labeled P, Q, and R represent the sets of those names. The rectangular region represents the universal set U. There are eight subregions, labeled 1 through 8 in the picture. Region 8 represents the subset $P \cap Q \cap R$; region 1 represents $U-(P \cup Q \cup R)$; region 2 represents the elements of $Q - (P \cup R)$; and so on.

Let's use the above Venn diagram to verify that the distributive rule, $P \cup (Q \cap R) = (P \cup Q) \cap (P \cup R)$, is valid. The idea is to replace the sets P, Q, and R by their corresponding sets of regions from the Venn diagram. Thus, Q is replaced by $\{2,5,6,8\}$, P is replaced by $\{4,6,7,8\}$, and R is replaced by $\{3,5,7,8\}$. Even though the sets P, Q, and R are arbitrary, perhaps even infinite, the distributive rule reduces to verifying the same rule for these simplified sets:

$$\{4,6,7,8\} \cup \big(\{2,5,6,8\} \cap \{3,5,7,8\}\big) = \big(\{4,6,7,8\} \cup \{2,5,6,8\}\big) \cap \big(\{4,6,7,8\} \cup \{3,5,7,8\}\big).$$

This identity is trivial to check directly: Both sides reduce to the set $\{4, 5, 6, 7, 8\}$.

This "Venn diagram" approach reduces a set identity that involves potentially infinitely many elements to subsets of a set of eight elements. It is fine for proofs and especially good for checking out "set identities" to see quickly if they are true or not. For example, is it true that $Q - (P \cap R) = Q - (P \cap Q \cap R)$? Checking the Venn diagram shows that both sides correspond to the set of regions $\{2, 5, 6\}$. The identity is true. You will get a chance to practice this technique in the exercises. $\square$

There are, of course, other ways to verify set identities. One way is called the *element method*:

Example 2 (The element method for proofs of set equations) To use that method, you simply translate the identity $X = Y$ into basic statements about what conditions a single element must satisfy to be (first) in the set on the left and then (second) in the set on the right. Thus, to show that $X = Y$, you assert that if $x \in X$ then blah, blah, blah (a bunch of words that make sense) implies that $x \in Y$. This shows that $X \subseteq Y$. Then, you reverse the argument and assert that if $y \in Y$ then blah, blah, blah (a bunch of words that make sense) implies that $y \in X$. This shows that $Y \subseteq X$. Thus $X = Y$.

Here is an example. Show, by the element method that, for all subsets P, Q, and R of U, $(P - Q) \cap (R - Q) = (P \cap R) - Q$.

(1) If $x \in (P - Q) \cap (R - Q)$ then (here comes the blah, blah, blah) x is in P but not in Q AND x is in R but not in Q.

(2) Thus x is in P **and** R, but x is not in Q.

(3) Thus x is in $(P \cap R) - Q$. This shows that $(P - Q) \cap (R - Q) \subseteq (P \cap R) - Q$. We leave it to you to use the element method to show the reverse, $(P - Q) \cap (R - Q) \supseteq (P \cap R) - Q$, and hence that $(P - Q) \cap (R - Q) = (P \cap R) - Q$. You should start your argument by saying, "Suppose $x \in (P \cap R) - Q$." $\square$

A different sort of element approach looks at each element of the universal set U and asks which sets contain it. The reult can be put in tabular form. When this is done, each row of the table corresponds to a region in the Venn diagram. The next example illustrates this *tabular method*.

Example 3 (The tabular method for proofs of set equations) We redo the identity of the previous example: $(P - Q) \cap (R - Q) = (P \cap R) - Q$. To do this we construct a table whose columns are labeled by various sets and whose entries answer the question "Is x in the set?" The first three columns in the following table are set up to allow all possible answers to the three questions "Is x in P?" "Is x in Q?" "Is x in R?" "Left" and "Right" refer to $(P - Q) \cap (R - Q)$ and $(P \cap R) - Q$, the two sides of the equation we want to prove. "Venn" refers to the region in the Venn diagram of Example 1. Normally that column would not be in the table, but we've inserted it so that you can see how each row corresponds to a Venn diagram region.

P	Q	R	$P - Q$	$R - Q$	Left	$P \cap R$	Right	Venn
No	No	No	No	No	No	No	No	1
No	No	Yes	No	Yes	No	No	No	3
No	Yes	No	No	No	No	No	No	2
No	Yes	Yes	No	No	No	No	No	5
Yes	No	No	Yes	No	No	No	No	4
Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	7
Yes	Yes	No	No	No	No	No	No	6
Yes	Yes	Yes	No	No	No	Yes	No	8

Since the answers are identical in the columns labeled "Left" and "Right," the identity is proved.

We can prove more from the table. For example,

$$\text{If } P \subseteq Q \cup R, \text{ then } P - Q = (P \cap R) - Q.$$

How does the table prove this? Because of the condition, the row that begins "Yes No No" is impossible. Therefore, we throw out that row and compare columns "$P - Q$" and "Right." $\square$

Another way to prove set identities is to use the basic algebraic identities of Theorem 1. This is called the *algebraic method*.

Example 4 (An algebraic proof) It is probably a good idea for you to label the steps with the appropriate rule (e.g., DeMorgan's rule, associative rule, distributive rule, etc.) the first few times you do such a proof. Therefore, we'll do that in this example. Mathematicians, however, would rarely bother to do it. A proof is accepted if others who know the basic rules of set theory can read it, understand it, and believe it is true.

Let's prove that $Q - (P \cap R) = Q - (P \cap Q \cap R)$. Here it is

$$\begin{aligned}
Q - (P \cap R) &= Q \cap (P \cap R)^c && \text{since } A - B = A \cap B^c \\
&= Q \cap (P^c \cup R^c) && \text{DeMorgan's rule} \\
&= (Q \cap P^c) \cup (Q \cap R^c) && \text{distributive rule} \\
&= (Q \cap P^c) \cup \emptyset \cup (Q \cap R^c) && \text{since } A \cup \emptyset = A \\
&= (Q \cap P^c) \cup (Q \cap Q^c) \cup (Q \cap R^c) && \text{since } Q \cap Q^c = \emptyset \\
&= Q \cap (P^c \cup Q^c \cup R^c) && \text{distributive rule} \\
&= Q \cap (P \cap Q \cap R)^c && \text{DeMorgan's rule} \\
&= Q - (P \cap Q \cap R) && \text{since } A - B = A \cap B^c
\end{aligned}$$

Some steps in this proof are baffling. For example, why did we introduce $\emptyset$ in the fourth line? We knew where we were going because we worked from both "ends" of the proof. In other words, we came up with a proof that moved both ends toward the middle and then rearranged the steps so that we could go from one end to the other. Unfortunately, proofs are often presented this way.

Here's another way to write the proof of $Q - (P \cap R) = Q - (P \cap Q \cap R)$ that shows more clearly how we got the proof. Note first that this identity is equivalent to showing that

$$Q \cap (P \cap R)^c = Q \cap (P \cap Q \cap R)^c$$

since $A - B = A \cap B^c$. This is equivalent, by DeMorgan's rules, to showing that

$$Q \cap (P^c \cup R^c) = Q \cap (P^c \cup Q^c \cup R^c).$$

But

$$Q \cap (P^c \cup Q^c \cup R^c) = Q \cap \Big((P^c \cup R^c) \cup Q^c\Big) = \Big(Q \cap (P^c \cup R^c)\Big) \cup \Big(Q \cap Q^c\Big) = Q \cap (P^c \cup R^c).$$

This latter identity follows from the fact that $Q \cap Q^c = \emptyset$ and $X \cup \emptyset = X$ for any set X. This completes the proof.

How should you write an algebraic proof? You can use whichever method you prefer. The first approach can be read mechanically because of the way it's laid out. However, if you use the first approach, you may sometimes need to use the second method for yourself first. $\square$

Ordering Sets

In computer programming, you will store and compute with sets of all sorts (sets of number, letters, geometric figures, addresses to arrays, pointers to structures, etc.). In almost all cases, you will work with these sets as lists (also called "linear orders") of some type where order does matter. The order matters in terms of the efficiency of your computations, not in terms of the rules of set theory.

In many cases, the linear ordering of the elements of a set is inherited from the universal set U. For example, the sets $A = \{1, 2, 3\}$ and $B = \{a, b, c\}$ inherit a natural linear ordering from the integers and the alphabet, respectively. But what about $C = \{?, >, <, \}$? There is no standard convention for C. You could use the ASCII code order $(<, >, ?)$, but if you do, some explanation should be given.

Example 5 (Lexicographic order) If you have decided on linear orders (i.e., listings) for a set X and a set Y, there is a commonly used and natural linear ordering for $X \times Y$ called *lexicographic order*. Suppose we list X and Y in some manner:

$$(x_1, x_2, \ldots, x_n) \quad \text{and} \quad (y_1, y_2, \ldots, y_m).$$

Given pairs $(a,b) \in X \times Y$ and $(c,d) \in X \times Y$, we say that (a,b) is *lexicographically less than or equal* to (c,d) if

(1) a is before c in the linear order on X or

(2) $a = c$ and b is equal to or before d in the linear order on Y.

For example, the lexicographic order for $\{1,2,7\} \times \{a,b\}$ is

$$\big\{(1,a),\ (1,b),\ (2,a),\ (2,b),\ (7,a),\ (7,b)\big\}.$$

Lexicographic order is called *lex order* for short.

Once you have ordered $X \times Y$ lexicographically, you can order $(X \times Y) \times Z$ by the same two rules (1) and (2) above, provided an order is specified on Z. You can use lex order on $X \times Y$ and the given linear order on Z. Likewise, you can apply (1) and (2) to $X \times (Y \times Z)$ using the given linear order on X and lex order on $Y \times Z$. The sets $(X \times Y) \times Z$ and $X \times (Y \times Z)$ are different sets – elements in the former have the form $((x,y),z)$ and those in the latter have the form $(x,(y,z))$. Imagine listing $(X \times Y) \times Z$ in lex order and stripping off the inner parentheses so that $((x,y),z)$ becomes (x,y,z). Now do the same with $X \times (Y \times Z)$. The two lists will contain the same elements in the same order. It's easy to see why the elements are the same, but it's not so easy to see why the orders are the same. Let's prove the orders are the same.

That means we have to prove

$$((x_1,y_1),z_1) \text{ precedes } ((x_2,y_2),z_2) \text{ if and only if } (x_1,(y_1,z_1)) \text{ precedes } (x_2,(y_2,z_2)).$$

It will make things easier if we write "$((x_1,y_1),z_1) < ((x_2,y_2),z_2)$" for "$((x_1,y_1),z_1)$ precedes $((x_2,y_2),z_2)$." By definition the definition of lex order, $((x_1,y_1),z_1) < ((x_2,y_2),z_2)$ means that $(x_1,y_1) < (x_2,y_2)$ or $(x_1,y_1) = (x_2,y_2)$ and $z_1 < z_2$. By the definition of lex order, the first case means that either $x_1 < x_2$ or $x_1 = x_2$ and $y_1 < y_2$. Note that $(x_1,y_1) = (x_2,y_2)$ means $x_1 = x_2$ and $y_1 = y_2$. Putting all this together, we have shown that $((x_1,y_1),z_1) < ((x_2,y_2),z_2)$ means that either

(1) $x_1 < x_2$ or

(2) $x_1 = x_2$ and $y_1 < y_2$ or

(3) $x_1 = x_2$ and $y_1 = y_2$ and $z_1 < z_2$.

In other words, look for the first position where (x_1,y_1,z_1) and (x_2,y_2,z_2) disagree and use that position to determine the order. We leave it to you to show that the same conditions describe $(x_1,(y_1,z_1)) < (x_2,(y_2,z_2))$. This completes the proof.

Since $(X \times Y) \times Z$ and $X \times (Y \times Z)$ have the same order when inner parentheses are dropped, one usually does that. We just write $X \times Y \times Z$ and call the elements (x,y,z).

When you think about what we've just done, you should be able to see that, if $S_1, S_2, \ldots, S_n$ are sets, we can write $S_1 \times S_2 \times \cdots \times S_n$, leaving out parentheses. The product $S_1 \times S_2 \times \cdots \times S_n$ is also written $\times_{i=1}^n S_i$.

Suppose $(s_1, s_2, \ldots, s_n)$ and $(t_1, t_2, \ldots, t_n)$ are in $S_1 \times S_2 \times \cdots \times S_n$. Which comes first? You should be able to see that we determine which precedes which as follows: By going left to right, find the first position where they disagree. Say position k is where this disagreement occurs. Use the order of s_k and t_k to determine the order of $(s_1, s_2, \ldots, s_n)$ and $(t_1, t_2, \ldots, t_n)$. This is the same order you use when you look things up in a dictionary. $\square$

Example 6 (Dictionary order on words or strings) The order of words in the dictionary is called "dictionary order." Lex order appears to be the same as dictionary order, but there is a problem with this. We've only defined lex order for n-tuples where n has some fixed value, but words in the dictionary have different length. Let's look at this more carefully.

Let S be a finite set. Let $S^k = \times^k S$ be the product of S with itself k times. The set S^0 is special and consists of one string ϵ called the *empty string*. Let $S^* = \cup_{k=0}^{\infty} S^k$. In words, S^* consists of all strings (words, vectors) of all possible lengths (including length zero) over S. Assume S is linearly ordered. We now define an order relation on S^* called *lexicographic order* or *dictionary order*, denoted by $\leq_L$, on S^*.

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

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

if either of the following two conditions hold:

(D1) $m \leq n$ and $a_i = b_i$ for $i = 1, \ldots, m$.

(D2) 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} is before b_{k+1} in the linear order on S.

Since $m, n > 0$, we have not discussed the empty string. Thus we need:

(D3) The empty string $\epsilon \leq_L x$ for any string x.

We have just defined dictionary order and also called it lex order. Is this the same as our previous definition of lex order? Yes because the two definitions of lex order agree when the strings have the same length.

We shall study this ordering on words carefully when we study order relations in general. For now we just give an example. Let $S = \{x, y\}$ with the ordering on S the alphabetic order. If $u = (x, x, y)$ and $v = (x, x, y, x)$, then $u \leq_L v$ by (D1). If $s = (x, x, y, x)$ and $t = (x, x, x, y)$, then $t \leq_L s$ by (D2). More examples will be given in the exercises. The standard English dictionary is an example where this linear order is applied to a subset of all words on the standard English alphabet (the words that have meaning in English).

A variation on this dictionary order is to order all words first by length and then by lex order. Thus, $u = (y, y, y)$ comes before $v = (x, x, x, x)$ because u has length three (three components) and v has length one. This order on S^* is called *length-first lex order* or *short lex order*. $\square$

Subsets of Sets

We use the notation $\mathcal{P}(A)$ to denote the set of all subsets of A and $\mathcal{P}_k(A)$ the set of all subsets of A of size (or cardinality) k. We call $\mathcal{P}(A)$ "the set of all subsets of A" or simply the *power set* of A. Let $C(n,k) = |\mathcal{P}_k(A)|$ denote the number of different k-subsets that can be formed from an n-set. The notation $\binom{n}{k}$ is also frequently used. These are called *binomial coefficients* and are read "n choose k." We now prove

Theorem 2 (Binomial coefficient formula) *The value of the binomial coefficient is*

$$\binom{n}{k} = C(n,k) = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!\,(n-k)!},$$

where $0! = 1$ and, for $j > 0$, $j!$ is the product of the first j integers. We read $j!$ as "j factorial".

Proof: Let A be a set of size n. The elements of $\mathcal{P}_k(A)$ are sets and are thus unordered. Generally speaking, unordered things are harder to count than ordered ones. Suppose, instead of a set of size k chosen from A, you wanted to construct an ordered list L of k elements from A (L is called a "k-list"). We could construct L in two stages.

- First choose an element of $S \in \mathcal{P}_k(A)$ (a subset of A with k elements). This can be done in $C(n,k)$ ways since $C(n,k) = |\mathcal{P}_k(A)|$.
- Next order S to obtain L. This ordering can be done in $k! = k(k-1)\cdots 1$ ways. Why? You have k choices for the element of S to appear first in the list L, $k-1$ choices for the next element, $k-2$ choices for the next element, etc.

From this two-stage process, we see that there are $C(n,k)\,k!$ ordered k-lists with no repeats. (The factor $C(n,k)$ is the number of ways to carry out the first stage and the factor $k!$ is the number of ways to carry out the second stage.)

Theorem 3 (Number of ordered lists) *The number of ordered k-lists L that can be made from and n-set A is*

- *n^k if repeats are allowed and*
- *$n(n-1)\cdots(n-k+1) = n!/(n-k)!$ if repeats are not allowed. One also uses the notation $(n)_k$ for these values. This is called the "falling factorial" and is read "n falling k".*

Why? With repeats allowed, there are n choices of elements in A for the first entry in the k-list L, n choices for the second entry, etc. If repeats are not allowed, there are n choices of elements in A for the first entry in the k-list L, $n-1$ choices for the second entry, etc.

Since we've counted the same thing (k-lists made from A) in two different ways, the two answers must be equal; that is, $C(n,k)\,k! = n!/(n-k)!$. Dividing by $k!$, we have the theorem. $\square$

In high school, you learned about "Pascal's Triangle" for computing binomial coefficients. We review this idea in the next example.

Example 7 (Binomial recursion) Let $X = \{x_1, \ldots, x_n\}$. We'll think of $C(n,k)$ as counting k-subsets of X. Imagine that we are going to construct a subset S of X with k elements. Either the element x_n is in our subset S or it is not. The cases where it is in the subset S are all formed by taking the various $(k-1)$-subsets of $X - \{x_n\}$ and adding x_n to them. By the definition of binomial coefficients, there are $\binom{n-1}{k-1}$ such subsets. The cases where it is not in the subset S are all formed by taking the various k-subsets of $X - \{x_n\}$. By the definition of binomial coefficients, there are $\binom{n-1}{k}$ such subsets. What we've done is describe how to build all k-subsets of X from certain subsets of $X - \{x_n\}$. Since this gives each subset exactly once,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k},$$

which can be written $C(n,k) = C(n-1,k-1) + C(n-1,k)$. This equation is called a *recursion* because it tells how to compute the function $C(n,k)$ from values of the function with smaller arguments. Here are the starting values together with the basic recursion:

$$\begin{aligned} &C(1,0) = C(1,1) = 1, \\ &C(1,k) = 0 \quad \text{for } k \neq 0,1 \quad \text{and} \\ &C(n,k) = C(n-1,k-1) + C(n-1,k) \quad \text{for } n > 1. \end{aligned}$$

Below we have made a table of values for $C(n,k)$.

$n \backslash k$	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1

C(n,k)

This tabular representation of $C(n,k)$ is called "Pascal's Triangle." $\square$

Definition 2 (Characteristic function) *Let U be the universal set and let $A \subseteq U$. The characteristic function of A, denoted χ_A is defined for each $x \in U$ by*

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

Thus the domain of χ_A is U and the range of χ_A is $\{0,1\}$.[2]

[2] *If you are not familiar with "domain" and "range", see the definition at the beginning*

Example 8 (Subsets as (0,1)-vectors) If A has n elements, listed $(a_1, a_2, \ldots, a_n)$, then you can specify any subset $X \subset A$ by a sequence $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ where $\epsilon_k = 0$ if the element $a_k \notin X$ and $\epsilon_k = 1$ if the element $a_k \in X$. The vector $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ is just the characteristic function of X since $\epsilon_k = \chi_X(a_k)$.

How many different subsets of A are there? We'll show that there are 2^n choices for $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ and thus $|\mathcal{P}(A)| = 2^n$. Why 2^n? There are clearly two choices for ϵ_1 and two choices for ϵ_2 and so forth. Thus there are $2 \times 2 \times \cdots = 2^n$ choices for $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$. □

Example 9 (Sets with sets as elements) Sets can have sets as elements. In the first exercise of this section, you will be asked such questions as "Is $\{1,2\} \in \{\{1,2\},\{3,4\}\}$?" or "Is $1 \in \{\{1\},\{2\},\{3\}\}$?" Easy stuff if you understand the definitions: You can see that the set $\{1,2\}$ is indeed an element of the set $\{\{1,2\},\{3,4\}\}$ because this latter set has just two elements, each of them a set of size two, one of which is $\{1,2\}$. You can also see that every element of $\{\{1\},\{2\},\{3\}\}$ is a set and that the number 1 is nowhere to be found as an element of this set.

You have already seen $\mathcal{P}(A)$, which is a set whose elements are sets, namely the subsets of A.

Another important class of sets with sets as elements are the *set partitions*. Some of the elementary aspects of set partitions fit into our present discussion. More advanced aspects of them will be discussed in Section 2. Here is a preview. Let $A = \{1, 2, \ldots, 15\}$. Consider the following set whose elements are themselves subsets of A.

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

This set is a subset of the power set $\mathcal{P}(A)$. But, it is a very special type of subset, called a *set partition* of A because it satisfies the three conditions:

(1) every element of α is nonempty,

(2) the union of the elements of α is A, and

(3) if you pick sets $X \in \alpha$ and $Y \in \alpha$, either $X = Y$ or $X \cap Y = \emptyset$.

Any collection of subsets of a set A satisfying (1), (2), and (3) is a set partition of A or simply a partition of A. Each element of α (which is, of course, a subset of A) is called a *block* of the partition α.

How many partitions are there of a set A? This is a tricky number to compute and there is no simple formula like $C(n,k) = \frac{n!}{k!(n-k)!}$ for it. We will discuss it in the Section 2. The number of partitions of a set of size n is denoted by B_n. These numbers are called *Bell numbers* after Eric Temple Bell. The first few Bell numbers are $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$.

We can *refine* the partition α by splitting blocks into smaller blocks. For example, we might split the block $\{6,8,10,15\}$ into two blocks, say $\{6,15\}$ and $\{8,10\}$, and also split the block $\{11,12,13,14\}$ into three blocks, say $\{13\}$, $\{14\}$, and $\{11,12\}$. The resulting partition is called a *refinement* of α and equals

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

---

of the next section.

Note that a refinement of a partition is another partition of the same set. We also consider a partition α to be a refinement of itself. We shall gain a deeper understanding of the notion of refinement when we study order relations. $\square$

Exercises for Section 1

1.1. Answer the following about the $\in$ and $\subseteq$ operators.

(a) Is $\{1,2\} \in \{\{1,2\},\{3,4\}\}$?

(b) Is $\{2\} \in \{1,2,3,4\}$?

(c) Is $\{3\} \in \{1,\{2\},\{3\}\}$

(d) Is $\{1,2\} \subseteq \{1,\ 2,\ \{1,2\},\ \{3,4\}\}$?

(e) Is $1 \in \{\{1\},\ \{2\},\ \{3\}\}$?

(f) Is $\{1,2,1\} \subseteq \{1,2\}$?

1.2. For each of the following, draw a Venn diagram.

(a) $A \subseteq B$, $C \subseteq B$, $A \cap C = \emptyset$

(b) $A \supseteq C$, $B \cap C = \emptyset$.

1.3. Let $A = \{w,x,y,z\}$ and $B = \{a,b\}$. Take the linear orders on A and B to be alphabetic order. List the elements in each of the following sets in lexicographic order.

(a) $A \times B$

(b) $B \times A$

(c) $A \times A$

(d) $B \times B$

1.4. Let $A = \{1,2,3\}$, $B = \{u,v\}$, and $C = \{m,n\}$. Take the linear order on A to be numeric and the linear orders on B and C to be alphabetic. List the elements in each of the following sets in lexicographic order.

(a) $A \times (B \times C)$ (use lex order on $B \times C$).

(b) $(A \times B) \times C$ (use lex order on $A \times B$).

(c) $A \times B \times C$.

1.5. Let $\Sigma = (\mathrm{x},\mathrm{y})$ be an alphabet. List each of the following sets of strings over this alphabet in the order indicated.

(a) All palindromes (strings that read the same forward and backward) of length less than or equal to 4. List them in dictionary order.

(b) All strings (words) that begin with x and have length less than four. List them in both dictionary and length-first lex order.

(c) List all strings of length four in lex order.

1.6. Each of the following statements about subsets of a set U is FALSE. Draw a Venn diagram to represent the situation being described. In each case case, show that the assertion is false by specializing the sets.

(a) For all A, B, and C, if $A \not\subseteq B$ and $B \not\subseteq C$ then $A \not\subseteq C$.

(b) For all sets A, B, and C, $(A \cup B) \cap C = A \cup (B \cap C)$.

(c) For all sets A, B, and C, $(A - B) \cap (C - B) = A - (B \cup C)$.

(d) For all A, B, and C, if $A \cap C \subseteq B \cap C$ and $A \cup C \subseteq B \cup C$ then $A = B$.

(e) For all A, B, and C, if $A \cup C = B \cup C$ then $A = B$.

(f) For all sets A, B, and C, $(A - B) - C = A - (B - C)$.

1.7. Prove each statement directly from the definitions.

(a) If A, B, and C are subsets of U, then $A \subseteq B$ and $A \subseteq C$ implies that $A \subseteq B \cap C$.

(b) If A, B, and C are subsets of U, then $A \subseteq B$ and $A \subseteq C$ implies that $A \subseteq B \cup C$.

1.8. Prove, using the definition of set equality, that for all sets A, B, and C, $(A - B) \cap (C - B) = (A \cap C) - B$.

1.9. Prove each statement by the method indicated.

(a) Prove using element arguments that if U is the universal set and A and B subsets of U, then $A \subseteq B$ implies that $U - A \supseteq U - B$ (alternative notation: $A \subseteq B$ implies $A^c \supseteq B^c$, or $A' \supseteq B'$)

(b) Prove, using element arguments and the definition of set inclusion, that for all A, B, and C, if $A \subseteq B$ then $A \cap C \subseteq B \cap C$.

(c) Prove, using (a), (b), and DeMorgan's law, that for all A, B, and C, if $A \subseteq B$ then $A \cup C \subseteq B \cup C$.

1.10. Prove each statement by the "element method."

(a) If A, B, and C are subsets of U, then $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

(b) If A, B, and C are subsets of U, then $A \times (B \cap C) = (A \times B) \cap (A \times C)$.

1.11. Prove each of the following identities from the basic algebraic rules for sets. You may want to use the fact that $D - E = D \cap E^c$.

(a) If A, B, and C are subsets of U, then $(A - B) - C = A - (B \cup C)$.

(b) If A, B, and C are subsets of U, then $(A - B) - C = (A - C) - B$.

(c) If A and B are subsets of U, then $(A - B) \cup (B - A) = (A \cup B) - (A \cap B)$.

1.12. Prove or give a counterexample. Use a Venn diagram argument for the proof. For the counterexample, use a Venn diagram or use set specialization.

(a) If A, B, and C are subsets of U, then $(A - C) \cap (B - C) \cap (A - B) = \emptyset$.

(b) If A and B are subsets of U and if $A \subseteq B$, then $A \cap (U - B) = \emptyset$.

(c) If A, B, and C are subsets of U, and if $A \subseteq B$, then $A \cap (U - (B \cap C)) = \emptyset$.

(d) If A, B, and C are subsets of U, and if $(B \cap C) \subseteq A$, then $(A - B) \cap (A - C) = \emptyset$.

(e) If A and B are subsets of U and if $A \cap B = \emptyset$, then $A \times B = \emptyset$.

1.13. Recall that the symmetric difference of sets A and B is $A \oplus B = (A - B) \cup (B - A)$. It is evident from the definition that $A \oplus B = B \oplus A$, the commutative law. Let U be the universal set. Prove each of the following properties either using a Venn diagram argument or algebraically or directly from the definition.

(a) $A \oplus (B \oplus C) = (A \oplus B) \oplus C$ (associative law for $\oplus$).

(b) $A \oplus \emptyset = A$.

(c) $A \oplus A^c = U$.

(d) $A \oplus A = \emptyset$.

(e) If $A \oplus C = B \oplus C$ then $A = B$.

1.14. Let A, B, and C be subsets of U. Prove or disprove using Venn diagrams.

(a) $A - B$ and $B - C$ are disjoint.

(b) $A - B$ and $C - B$ are disjoint.

(c) $A - (B \cup C)$ and $B - (A \cup C)$ are disjoint.

(d) $A - (B \cap C)$ and $B - (A \cap C)$ are disjoint.

1.15. Which of the following are partitions of $\{1, 2, \ldots, 8\}$? Explain your answers.

(a) $\big\{\{1,3,5\}, \{1,2,6\}, \{4,7,8\}\big\}$

(b) $\big\{\{1,3,5\}, \{2,6,7\}, \{4,8\}\big\}$

(c) $\big\{\{1,3,5\}, \{2,6\}, \{2,6\}, \{4,7,8\}\big\}$

(d) $\big\{\{1,5\}, \{2,6\}, \{4,8\}\big\}$

1.16. How many refinements are there of the partition $\big\{\{1,3,5\}, \{2,6\}, \{4,7,8,9\}\big\}$? Explain.

1.17. Suppose S and T are sets with $S \cap T = \emptyset$. Suppose σ is a partition of S and τ is a partition of T.

(a) Prove that $\sigma \cup \tau$ is a partition of $S \cup T$.

(b) If σ has n_σ refinements and τ has n_τ refinements, how many refinements does $\sigma \cup \tau$ have? Explain.

1.18. Use the characteristic function format to list the power set of the following sets. That is, describe each element of the power set as a vector of zeroes and ones.

(a) $\{1, 2, 3\}$

(b) $X \times Y$ where $X = \{a, b\}$ and $Y = \{x, y\}$.

1.19. Find the following power sets:

(a) $\mathcal{P}(\emptyset)$

(b) $\mathcal{P}(\mathcal{P}(\emptyset))$

(c) $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$

1.20. Compare the following pairs of sets. Can they be equal? Is one a subset of the other? Can they have the same size (number of elements)?

(a) $\mathcal{P}(A \cup B)$ and $\mathcal{P}(A) \cup \mathcal{P}(B)$

(b) $\mathcal{P}(A \cap B)$ and $\mathcal{P}(A) \cap \mathcal{P}(B)$

(c) $\mathcal{P}(A \times B)$ and $\mathcal{P}(A) \times \mathcal{P}(B)$

1.21. Let $S = \{1, 2, \ldots, n\}$. Let S_1 be the set of all subsets of S that contain 1. Let T_1 denote the set of all subsets of S that don't contain 1. Prove $|T_1| = |S_1| = 2^{(n-1)}$.

Section 2: Functions

Functions, such as linear functions, polynomial functions, trigonometric functions, exponential functions, and logarithmic functions are familiar to all students who have had mathematics in high school. For discrete mathematics, we need to understand functions at a basic set theoretic level. We begin with a familiar definition.

Definition 3 (Function) *If A and B are sets, a function from A to B is a rule that tells us how to find a unique $b \in B$ for each $a \in A$. We write $f(a) = b$ and say that f maps a to b. We also say the value of f at a is b.*

We write $f : A \to B$ to indicate that f is a function from A to B. We call the set A the domain of f and the set B the range or, equivalently, codomain of f.

To specify a function completely you must give its domain, range and rule.

If $X \subseteq A$, then $f(X) = \{f(x) \mid x \in X\}$. In particular $f(\emptyset) = \emptyset$ and $f(A)$ is called the image of f.

Some people define "range" to be the values that the function *actually* takes on. Most people call that the *image*.

In high school, you dealt with functions whose ranges were $\mathbb{R}$ and whose domains were contained in $\mathbb{R}$; for example, $f(x) = 1/(x^2 - 1)$ is a function from $\mathbb{R} - \{-1, 1\}$ to $\mathbb{R}$. If you have had some calculus, you also studied functions of functions! The derivative is a function whose domain is all differentiable functions and whose range is all functions. If we wanted to use functional notation we could write $D(f)$ to indicate the function that the derivative associates with f.

The set of all functions from A to B is written B^A. One reason for this notation, as we shall see below, is that $|B^A| = |B|^{|A|}$. Thus $f : A \to B$ and $f \in B^A$ say the same thing.

To avoid the cumbersome notation $\{1, 2, 3, \ldots, n\}$, we will often use $\underline{n}$ instead.

Example 10 (Functions as relations) There is a fundamental set-theoretic way of defining functions. Let A and B be sets. A *relation from A to B* is a subset of $A \times B$. For example, if $A = \underline{3} = \{1, 2, 3\}$ and $B = \underline{4}$, then $R = \{(1, 4), (1, 2), (3, 3), (2, 3)\}$ is a relation from A to B. To specify a relation, you must define three sets: A, B and R.

If the relation R satisfies the condition $\forall\, x \in A\ \exists!\, y \in B,\ (x, y) \in R$, then the relation R is called a *functional relation*. We used some shorthand notation here that is worth remembering:

$\forall$	means	"for all"
$\exists$	means	"for some" or "there exists"
$\exists!$	means	"for exactly one"

If you think about Definition 3, you will realize that a "functional relation" is just one possible way of giving all of the information required to specify a function.

Given any relation $R \subseteq A \times B$, the inverse relation R^{-1} from B to A is $\{(y, x) : (x, y) \in R\}$. For $R = \{(1, 4), (1, 2), (3, 3), (2, 3)\}$, $A = \underline{3}$ and $B = \underline{4}$, the inverse relation is $R^{-1} = \{(4, 1), (2, 1), (3, 3), (3, 2)\}$. Note that neither R nor R^{-1} is a functional relation in this example. You should make sure that you understand why this statement is true. (Hint: R fails the "$\exists!$" test and R^{-1} fails the "$\forall$" part of the definition of a functional relation.) Note also that if R and R^{-1} are functional then $|A| = |B|$. In algebra or calculus, when you draw a graph of a real-valued function $f : \mathbf{D} \to \mathbb{R}$ (such as $f(x) = x^3$, $f(x) = x/(1 - x)$, etc.), you are attempting a pictorial representation of the set $\{(x, f(x)) : x \in \mathbf{D} \subseteq \mathbb{R}\}$, which is the subset of $\mathbf{D} \times \mathbb{R}$. This subset is a "functional relation from $\mathbf{D}$ to $\mathbb{R}$."

In our notation, we would write $(a, b) \in R$ to indicate that the pair (a, b) is in the relation R from A to B. People also use the notation $a\ R\ b$ to indicate this. For example, the "less than" relation $\{(a, b) \mid a < b\}$ is written $a < b$. $\square$

In many cases in discrete mathematics, we are concerned with functions whose domain is finite. Special notation is used for specifying such functions.

Definition 4 (One-line notation) *Let A be a finite ordered set with elements ordered $(a_1, a_2, \ldots, a_{|A|})$. Let B be any set. A function $f : A \to B$ can be written in one-line notation as $f = (f(a_1), f(a_2), \ldots, f(a_{|A|}))$. Thus the values of the function are written as* *list, which is also called a vector or a string. In other words the function f assigns to a_k the k^{th} element of the list $(f(a_1), f(a_2), \ldots, f(a_{|A|}))$ for each value of k from 1 to $|A|$.*

It follows from the definition that we can think of function as an element of $B^{|A|} = B \times B \times \cdots \times B$, where there are $|A|$ copies of B. This is another reason for the notation B^A for all functions from A to B. Do you see why we don't use $B^{|A|}$ instead? No, it's not because B^A is easier to write. It's because $B^{|A|}$ does not specify the domain A. Instead, only its size $|A|$ is given.

Example 11 (Using the notation) To get a feeling for the notation used to specify a function, it may be helpful to imagine that you have an envelope or box that contains a function. In other words, this envelope contains all the information needed to completely describe the function. Think about what you're going to see when you open the envelope.

You might see

$$P = \{a, b, c\}, \qquad g : P \to \underline{4}, \qquad g(a) = 3, \quad g(b) = 1 \quad \text{and} \quad g(c) = 4.$$

This tells you that the name of the function is g, the domain of g is P, which is $\{a, b, c\}$, and the range of g is $\underline{4} = \{1, 2, 3, 4\}$. It also tells you the values in $\underline{4}$ that g assigns to each of the values in its domain. Someone else may have put

$$g \in \underline{4}^{\{a,b,c\}}, \qquad \text{ordering: } a, b, c, \qquad g = (3, 1, 4).$$

in the envelope instead. This describes the same function. It doesn't give a name for the domain, but that's okay since all we need to know is what's in the domain. On the other hand, it gives an order on the domain so that the function can be given in one-line form. Since the domain is ordered a, b, c and since $g = (3, 1, 4)$, by the definition of one-line notation $g(a) = 3$, $g(b) = 1$ and $g(c) = 4$. Can you describe other possible envelopes for the same function?

What if the envelope contained only $g = (3, 1, 4)$? If you think you have been given the one-line notation for g, you are mistaken. You *must* know the *ordered* domain of g before you can interpret $g = (3, 1, 4)$. Here we don't even know the domain as a set (or the range). The domain might by $\{a, b, c\}$, or $\{<, >, ?\}$, or any other 3-set.

What if the envelope contained

$$\text{the domain of } g \text{ is } \{a, b, c\}, \qquad \text{ordering: } a, b, c, \qquad g = (3, 1, 4)?$$

We haven't specified the range of g, but is it necessary since we know the values of the function? Our definition included the requirement that the range be specified, so this is not a complete definition. Some definitions of a function do not require that the range be specified. For such definitions, this would be a complete specification of the function g. □

Example 12 (Counting functions) Think about specifying $f : A \to B$ in one-line notation: $(f(a_1), f(a_2), \ldots, f(a_{|A|}))$. There are $|B|$ ways to choose $f(a_1)$, $|B|$ ways to choose $f(a_2)$, etc., and finally $|B|$ ways to choose $f(a_{|A|})$. This means that the cardinality of the set of all functions $f : A \to B$ is $|B|^{|A|}$. In other words, $\left|B^A\right| = |B|^{|A|}$.

We can represent a subset S of A by a unique function $f : A \to \underline{2}$ where

$$f(x) = \begin{cases} 1, & \text{if } x \notin S, \\ 2, & \text{if } x \in S. \end{cases}$$

This proves that there are $2^{|S|}$ such subsets. We proved this result in Example 9. You should verify that this is essentially the same proof that was given there.

We can represent a list of k elements of a set S with repetition allowed by a unique function $f : \underline{k} \to S$. In this representation, the list corresponds to the function written in one-line notation. (Recall that the ordering on $\underline{k}$ is the numerical ordering.) This proves that there are exactly $|S|^k$ such lists. $\square$

Definition 5 (Types of functions) *Let $f : A \to B$ be a function. If for every $b \in B$ there is an $a \in A$ such that $f(a) = b$, then f is called a surjection (or an onto function). Another way to describe a surjection is to say that it takes on each value in its range at least once.*

If $f(x) = f(y)$ implies $x = y$, then f is called an injection (or a one-to-one function). Another way to describe an injection is to say that it takes on each value in its range at most once.

If f is both an injection and a surjection, it is a called a bijection. The bijections of A^A are called the permutations of A. The set of permutations on a set A is denoted in various ways. Two notations are $\mathrm{PER}(A)$ *and* $\mathcal{S}(A)$.

If $f : A \to B$ is a bijection, we may talk about the inverse bijection of f, written f^{-1}, which reverses what f does. Thus $f^{-1} : B \to A$ and $f^{-1}(b)$ is that unique $a \in A$ such that $f(a) = b$.

Note that $f(f^{-1}(b)) = b$ and $f^{-1}(f(a)) = a$. Do not confuse f^{-1} with $1/f$. For example, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^3 + 1$, then $1/f(x) = 1/(x^3 + 1)$ and $f^{-1}(x) = (x - 1)^{1/3}$.

Example 13 (Surjections, injections and bijections as lists) Lists provide another fundamental way to think about the various types of functions we've just defined. We'll illustrate this with some examples.

Let $A = \underline{4}$, $B = \{a, b, c, d, e\}$ and $f = (d, c, d, a)$ describe the function f in one-line notation. Since the list (d, c, d, a) contains d twice, f is not an injection. The function $(b, d, c, e) \in B^A$ is an injection since there are no repeats in the list of values taken on by the function. The 4-lists without repeats that can be formed from B correspond to the injections from $\underline{4}$ to B. In general, the injections in $S^{\underline{k}}$ correspond to k-lists without repeats whose elements are taken from S.

With the same f as in the previous paragraph, note that the value b is not taken on by f. Thus f is not a surjection. (We could have said e is not taken on, instead.)

Now let $A = \underline{4}$, $B = \{x, y, z\}$ and $g = (x, y, x, z)$. Since every element of B appears at least once in the list of values taken on, g is a surjection.

Finally, let $A = B = \underline{4}$ and $h = (3, 1, 4, 2)$. The function is both an injection and a surjection. Hence, it is a bijection. Since the domain and range are the same and f is a bijection, it is a permutation of $A = \underline{4}$. The list $(3, 1, 4, 2)$ is a rearrangement (a permutation) of the ordered listing $(1, 2, 3, 4)$ of A. That's why we call h a permutation. The inverse of h is $(2, 4, 1, 3)$. □

Example 14 (Encryption) Suppose we want to send data (a text message, a JPEG file, etc.) to someone and want to be sure no one else can read the data. Then we use *encryption*. We can describe encryption as a function $f : D \to R$ where D is the set of possible messages. Of course, there are a huge number of *possible* messages, so what do we do? We can break the message into pieces. For example, we could break an ordinary text message into pieces with one character (with the space as a character) per piece. Then apply a function f to each piece. Here's a simple example: If x is a letter, $f(x)$ is the next letter in the alphabet with $Z \to A$ and $f(\text{space}) = \text{space}$. Then we would encrypt "HELLO THERE" as "IFMMP UIFSF." This is too simple for encryption.

What can we do? Let S be the set of symbols that we are using (A to Z and space in the previous paragraph). We could choose a more complicated function $f : S \to S$ than our simple function. What properties should it have?

- It must have an inverse so that we can decrypt.
- The encryption and decryption must be quick on a computer.
- It must be hard for someone else to figure out f^{-1}.

Since $f : S \to S$ and it has an inverse, f must be a bijection (in fact, a permutation of S). How can we make f hard to figure out? That is a problem in the design of encryption systems. One key ingredient is to make S large. For example, in systems like PGP (Pretty Good Privacy) and DES (Data Encryption Standard) S consists of all n-long vectors of zeroes and ones, typically with $n = 64$. In this case $|S| = 2^{64} \approx 10^{19}$, which is quite large. □

Example 15 (Hashing) *Hashing* is a procedure for mapping a large space into a smaller one. For example, a hash function h may have as its domain all sequences of zeroes and ones of all possible lengths. It's range might be all n-long sequences of zeroes and ones for some n. There are some publicly available hash functions h that seem to be good.

Why would we want such a function? Suppose we want to be sure no one changes a document that is stored in a computer. We could apply h to the document and then save $h(\text{document})$. By giving $h(\text{document})$ to people, they could later check to see if the document had been changed — if the function h is well chosen it would be hard to change the document without changing the value of h, even if you know how to compute h. Suppose you email a document to a friend, but you're concerned that someone may intercept the email and change the document. You can call up your friend and tell him $h(\text{document})$ so that he can check it.

Another use for a hash function is storing data. Suppose we have an n-long array in which we want to store information about students at the university. We want a hash function that maps student ID numbers into $\{1, 2, \ldots n\}$. Then $h(\text{ID})$ tells us which array position to use. Of course two student ID numbers may hash to the same value (array position). There are methods for dealing with such conflicts. $\square$

Example 16 (Two-line notation) Since one-line notation is a simple, brief way to specify functions, it is used frequently. If the domain is not a set of numbers, the notation is poor because we must first pause and order the domain. There are other ways to write functions which overcome this problem. For example, we could write $f(a) = 4,\ f(b) = 3,\ f(c) = 4$ and $f(d) = 1$. This could be shortened up somewhat to $a \to 4$, $b \to 3$, $c \to 4$ and $d \to 1$. By turning each of these sideways, we can shorten it even more: $\begin{pmatrix} a & b & c & d \\ 4 & 3 & 4 & 1 \end{pmatrix}$. For obvious reasons, this is called *two-line notation*. Since x always appears directly over $f(x)$, there is no need to order the domain; in fact, we need not even specify the domain separately since it is given by the top line. If the function is a bijection, its inverse is obtained by interchanging the top and bottom lines.

The arrows we introduced in the last paragraph can be used to help visualize different properties of functions. Imagine that you've listed the elements of the domain A in one column and the elements of the range B in another column to the right of the domain. Draw an arrow from a to b if $f(a) = b$. Thus the heads of arrows are on elements of B and the tails are on elements of A. Since f is a function, no two arrows have the same tail. If f is an injection, no two arrows have the same head. If f is a surjection, every element of B is on the head of some arrow. You should be able to describe the situation when f is a bijection. $\square$

Example 17 (Compositions of functions) Suppose that f and g are two functions such that the values f takes on are contained in the domain of g. We can write this as $f : A \to B$ and $g : C \to D$ where $f(a) \in C$ for all $a \in A$. We define the *composition* of g and f, written $gf : A \to D$ by $(gf)(x) = g(f(x))$ for all $x \in A$. The notation $g \circ f$ is also used to denote composition. Suppose that f and g are given in two-line notation by

$$f = \begin{pmatrix} p & q & r & s \\ P & R & T & U \end{pmatrix} \qquad g = \begin{pmatrix} P & Q & R & S & T & U & V \\ 1 & 3 & 5 & 2 & 4 & 6 & 7 \end{pmatrix}.$$

Then $gf = \begin{pmatrix} p & q & r & s \\ 1 & 5 & 4 & 6 \end{pmatrix}$.

Suppose $f : A \to B$, $g : B \to C$, and $h : C \to D$. We can form the compositions $g \circ f$ and $h \circ g$; however, we cannot form the composition $h \circ f$ unless C contains $f(x)$ for all $x \in A$. We can also form the compositions of all three functions, namely $h \circ (g \circ f)$ and $(h \circ g) \circ f$. These two compositions are equal — that's the "associative law" for composition of functions. How is it proved? Here's an algebraic proof that uses nothing more than the definition of $\circ$ at each step: For all $x \in A$

$$\big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h\big(g(f(x))\big) = (h \circ g)(f(x)) = \big((h \circ g) \circ f\big)(x).$$

Let A be a set. Suppose that $f, g \in \mathcal{S}(A)$.; that is, f and g are permutations of a set A. Recall that a permutation is a bijection from a set to itself and so it makes sense to talk about f^{-1} and fg. We claim that fg and f^{-1} are also permutations of A. This is easy to see if you write the permutations in two-line form and note that the second line is a rearrangement of the first if and only if the function is a permutation.

Again suppose that $f \in \mathcal{S}(A)$. Instead of $f \circ f$ or ff we write f^2. Note that $f^2(x)$ is not $(f(x))^2$. (In fact, if multiplication is not defined in A, $(f(x))^2$ has no meaning.) We could compose three copies of f. The result is written f^3. In general, we can compose k copies of f to obtain f^k. A cautious reader may be concerned that $f \circ (f \circ f)$ may not be the same as $(f \circ f) \circ f$. By the associative law for $\circ$, they're equal. In fact, $f^{k+m} = f^k \circ f^m$ for all nonnegative integers k and m, where f^0 is defined by $f^0(x) = x$ for all x in the domain. This is true even if k or m or both are negative. $\square$

Example 18 (Composing permutations) Let's carry out some calculations for practice. Let f and g be the permutations

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix} \qquad g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}.$$

To compute fg, we must calculate $fg(x)$ for all x. This can be done fairly easily from the two-line form: For example, $(fg)(1)$ can be found by noting that the image of 1 under g is 2 and the image of 2 under f is 1. Thus $(fg)(1) = 1$. You should be able to verify that

$$fg = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 3 & 2 \end{pmatrix} \qquad gf = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} \neq fg.$$

Thus, $f \circ g = g \circ f$ (commutative law) is *not a law* for permutations.

It is easy to get the inverse, simply interchange the two lines. Thus

$$f^{-1} = \begin{pmatrix} 2 & 1 & 4 & 5 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} \quad \text{which is the same as } f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix},$$

since the order of the columns in two-line form does not matter.

Let's compute some powers:

$$f^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{pmatrix} \quad f^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix} \qquad g^5 = f^6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

We computed f^6 using $f^6 = f^3 \circ f^3$. That was a bit tedious. Now imagine if you wanted to compute f^{100}. Cycle notation is an easy way to do that. $\square$

Let f be a permutation of the set A and let $n = |A|$. If $x \in A$, we can look at the sequence $x, f(x), f(f(x)), \ldots, f^k(x), \ldots$, which is often written as $x \to f(x) \to f(f(x)) \to \cdots \to f^k(x) \to \ldots$. Using the fact that $f^0(x) = x$, we can write the sequence as $f^0(x) \to f^1(x) \to f^2(x) \to \cdots$. Since the range of f has n elements, this sequence will contain a repeated element in the first $n+1$ entries. Suppose that $f^s(x)$ is the first sequence entry that is ever repeated and that $f^p(x)$ is the first time that it is repeated.

We claim that $s = 0$. If $s > 0$, apply f^{-1} to both sides of this equality to obtain $f^{s-1}(x) = f^{p-1}(x)$, contradicting the fact that s was chosen as small as possible. Thus, in fact, $s = 0$.

It follows that the sequence cycles through a pattern of length p forever since $f^{p+1}(x) = f(f^p(x)) = f(x)$, $f^{p+2}(x) = f^2(f^p(x)) = f^2(x)$, and so on. We call $(x, f(x), \ldots, f^{p-1}(x))$ the *cycle* containing x and call p the *length of the cycle*. If a cycle has length p, we call it a p-cycle. Cyclic shifts of a cycle are considered the same; for example, if (1,2,6,3) is the cycle containing 1 (as well as 2, 3 and 6), then (2,6,3,1), (6,3,1,2) and (3,1,2,6) are other ways of writing the cycle.

Suppose $(x_1, x_2, \ldots, x_p)$ is a cycle of f and that $y_1 \in A$ is not in that cycle. We can form the cycle containing y_1: $(y_1, y_2, \ldots, y_q)$. None of the y_k is in the cycle $(x_1, \ldots, x_p)$. Why? If it were, we could continue in the cycle and eventually reach y_1. Written out algebraically: If $y_k = x_j$ for some k and j, then $y_1 = f^{q-k-1}(y_k) = f^{q-k-1}(x_j)$ and the right side is in the cycle $(x_1, \ldots, x_p)$. We have proved.

Theorem 4 (Cycle form of a permutation) *Let f be a permutation of the finite set A be a finite set. Every element of A belongs to a cycle of f. Two cycles are either the same or have no elements in common.*

Example 19 (Using cycle notation) Consider the permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 8 & 1 & 5 & 9 & 3 & 7 & 6 \end{pmatrix}.$$

Since $1 \to 2 \to 4 \to 1$, the cycle containing 1 is (1,2,4). We could equally well write it (2,4,1) or (4,1,2); however, (1,4,2) is different since it corresponds to $1 \to 4 \to 2 \to 1$. The usual convention is to list the cycle starting with its smallest element. The cycles of f are (1,2,4), (3,8,7), (5) and (6,9). We write f in cycle form as

$$f = (1, 2, 4)\ (3, 8, 7)\ (5)\ (6, 9).$$

The order in which the cycles are written doesn't matter, so we have

$$f = (5)\ (6, 9)\ (1, 2, 4)\ (3, 8, 7) \quad \text{and} \quad f = (4, 1, 2)\ (5)\ (6, 9)\ (7, 3, 8),$$

and lots of other equivalent forms. It is common practice to omit the cycles of length one and write $f = (1, 2, 4)(3, 8, 7)(6, 9)$. The inverse of f is obtained by reading the cycles backwards because $f^{-1}(x)$ is the lefthand neighbor of x in a cycle. Thus

$$f^{-1} = (4, 2, 1)(7, 8, 3)(9, 6) = (1, 4, 2)(3, 7, 8)(6, 9).$$

To compute $f(x)$, we simply take one step to the right from x in its cycle. We just saw that $f^{-1}(x)$ is computed by taking one step to the left. You may be able to guess at this point that $f^k(x)$ is computed by taking k steps to the right, with the rule that a negative step to the right is the same as a positive step to the left. This makes it easy to compute powers.

When $f = (1,2,4)(3,8,7)(5)(6,9)$, what is f^{100}? Imagine starting at 1. After 3 steps to the right we're back at 1. Do this 33 times so that after $3 \times 33 = 99$ steps to the right we're back at 1. One more step takes us to 100 steps and so $f^{100}(1) = 2$. You should be able to figure out the rest: $f^{100} = (1,2,4)(3,8,7)(5)(6)(9)$. $\square$

We next take a close look at the notions of *image* and *coimage* of a function. Again, let $f : A \to B$ be a function. The *image* of f is the set of values f actually takes on: $\mathrm{Image}(f) = \{\, f(a) \mid a \in A \,\}$. The definition of a surjection can be rewritten $\mathrm{Image}(f) = B$.

For each $b \in B$, the *inverse image* of b, written $f^{-1}(b)$ is the set of those elements in A whose image is b; i.e.,

$$f^{-1}(b) = \{\, a \mid a \in A \text{ and } f(a) = b \,\}.$$

This extends our earlier definition of f^{-1} from bijections to all functions; however, such an f^{-1} can't be thought of as a function from B to A unless f is a bijection because it will not give a unique $a \in A$ for each $b \in B$. (There is a slight abuse of notation here: If $f : A \to B$ is a bijection, our new notation is $f^{-1}(b) = \{a\}$ and our old notation is $f^{-1}(b) = a$.)

Definition 6 (Coimage) *Let $f : A \to B$ be a function. The collection of nonempty inverse images of elements of B is called the coimage of f. In set-theoretic terms,*

$$\mathrm{Coimage(f)} = \{f^{-1}(b) \mid b \in B,\ f^{-1}(b) \neq \emptyset\} = \{f^{-1}(b) \mid b \in \mathrm{Image}(f)\}.$$

To describe the structure of coimages, we need to recall that a partition of a set is an unordered collection of nonempty subsets of B such that each element of B appears in exactly one subset. Each subset is called a block of the partition.

Theorem 5 (Structure of coimage) *Suppose $f : A \to B$. The coimage of f is the partition of A whose blocks are the maximal subsets of A on which f is constant.*

Wait — before we give a proof we need to understand what we just said. Let's look at an example. If $f \in \{a,b,c\}^{\underline{5}}$ is given in one-line form as (a,c,a,a,c), then

$$\mathrm{Coimage}(f) = \{f^{-1}(a), f^{-1}(c)\} = \big\{\{1,3,4\},\{2,5\}\big\},$$

f is a on $\{1,3,4\}$ and is c on $\{2,5\}$. Now let's prove the theorem.

Proof: If $x \in A$, let $y = f(x)$. Then $x \in f^{-1}(y)$ and so the union of the nonempty inverse images contains A. Clearly it does not contain anything which is not in A. If $y_1 \neq y_2$, then we cannot have $x \in f^{-1}(y_1)$ and $x \in f^{-1}(y_2)$ because this would imply $f(x) = y_1$ and $f(x) = y_2$, a contradiction of the definition of a function. Thus $\mathrm{Coimage}(f)$ is a partition of A. Clearly x_1 and x_2 belong to the same block if and only if $f(x_1) = f(x_2)$. Hence a block is a maximal set on which f is constant. $\square$

Since $\mathrm{Coimage}(f)$ is a partition of the domain A, we need to review the basic combinatorial properties of partitions.

Example 20 (Set partitions) The 15 partitions of $\{1,2,3,4\}$, classified by number of blocks, are

1 block:	$\{1,2,3,4\}$			
2 blocks:	$\{\{1,2,3\},\{4\}\}$	$\{\{1,2,4\},\{3\}\}$	$\{\{1,2\},\{3,4\}\}$	$\{\{1,3,4\},\{2\}\}$
	$\{\{1,3\},\{2,4\}\}$	$\{\{1,4\},\{2,3\}\}$	$\{\{1\},\{2,3,4\}\}$	
3 blocks:	$\{\{1,2\},\{3\},\{4\}\}$	$\{\{1,3\},\{2\},\{4\}\}$	$\{\{1,4\},\{2\},\{3\}\}$	$\{\{1\},\{2,3\},\{4\}\}$
	$\{\{1\},\{2,4\},\{3\}\}$	$\{\{1\},\{2\},\{3,4\}\}$		
4 blocks:	$\{\{1\},\{2\},\{3\},\{4\}\}$			

Let $S(n,k)$ be the number of partitions of an n-set having exactly k blocks. These are called *Stirling numbers of the second kind*. Do not confuse $S(n,k)$ with $C(n,k) = \binom{n}{k}$. In both cases we have an n-set. For $C(n,k)$ we want to *choose a subset* containing k elements and for $S(n,k)$ we want to *partition the set* into k blocks.

What is the value of $S(n,k)$? Let's try to get a recursion. How can we build partitions of $\{1,2,\ldots,n\}$ with k blocks out of smaller cases? If we take partitions of $\{1,2,\ldots,n-1\}$ with $k-1$ blocks, we can simply add the block $\{n\}$. If we take partitions of $\{1,2,\ldots,n-1\}$ with k blocks, we can add the element n to one of the k blocks. You should convince yourself that all k block partitions of $\{1,2,\ldots,n\}$ arise in exactly one way when we do this. This gives us a recursion for $S(n,k)$. Putting n in a block by itself contributes $S(n-1,k-1)$. Putting n in a block with other elements contributes $S(n-1,k) \times k$ Thus,

$$S(n,k) = S(n-1,k-1) + k\,S(n-1,k).$$

Below is the tabular form for $S(n,k)$ analogous to the similar tabular form for $C(n,k)$.

$n \backslash k$	1	2	3	4	5	6	7
1	1						
2	1	1		$S(n,k)$			
3	1	3	1				
4	1	7	6	1			
5	1	15	25	10	1		
6	1	31	90	65	15	1	
7	1	--	--	--	--	--	1

Notice that the starting conditions for this table are that $S(n,1) = 1$ for all $n \geq 1$ and $S(n,n) = 1$ for all $n \geq 1$. The values for $n = 7$ are omitted from the table. You should fill them in to test your understanding of this computational process. For each n, the total number of partitions of a set of size n is equal to the sum $S(n,1) + S(n,2) + \cdots + S(n,n)$. These numbers, gotten by summing the entries in the rows of the above table, are the Bell numbers, $B(n)$, that we discussed in Section 1. $\square$

Example 21 (Counting functions by image size) Suppose A and B are sets. Let $|A| = m$ and $|B| = n$. Suppose $k \leq m$ and $k \leq n$. A basic question about functions $f : A \to B$ is the following:

$$\text{Let } S = \{f \mid f : A \to B,\ |\text{Image}(f)| = k\}. \text{ Find } |S|.$$

In other words, there are exactly k blocks in the coimage of f. This question clearly involves the Stirling numbers. In fact, the answer is $|S| = \binom{n}{k} S(m,k)\, k!$. The idea is to choose the image of the function in $\binom{n}{k}$ ways, then choose the coimage of the function in $S(m,k)$ ways and then put them together in $k!$ ways. You will get a chance to fill in the details in the last two exercises below. Here is an example. Suppose we take $|A| = 4$, $|B| = 5$, and $k = 3$. We get $|S| = \binom{5}{3} S(4,3)\, 3! = 10 \times 6 \times 6 = 360$.

Let's look at some special cases.

- If $k = |B|$, then we are counting surjections. Why? We are given $|\text{Image}(f)| = |B|$. Thus every element in B must be in Image(f).
- If $k = |A|$, then we are counting injections. Why? Suppose f is not an injection, say $f(a) = f(b)$ for some $a \neq b$. Then f can take on at most $|A| - 1$ different values. But $k = |A|$ says that $|\text{Image}(f)| = |A|$, a contradiction. $\square$

Exercises for Section 2

2.1. Let R be the relation on $\mathbb{R} \times \mathbb{R}$, the Cartesian plane, defined by xRy if $y = x^2$. Sketch a picture that represents the set R in $\mathbb{R} \times \mathbb{R}$.

2.2. Let R be a relation from the power set $\mathcal{P}(X)$ to itself, where $X = \{1, 2, 3, 4\}$, defined by $A\,R\,B$ if $A \cap B \neq \emptyset$.

(a) Is $A\,R\,A$ for all $A \in \mathcal{P}(X)$?

(b) For any $A, B \in \mathcal{P}(X)$, if $A\,R\,B$ is $B\,R\,A$?

(c) For any $A, B \in \mathcal{P}(X)$, if $A\,R\,B$ and $B\,R\,C$, is $A\,R\,C$?

2.3. In each case, draw the "directed graph diagram" of the given relation (label points in your diagram with the elements of S, put an arrow from x to y if and only if (x, y) belongs to the relation).

(a) $S = \{(a, b), (a, c), (b, c), (d, d)\}$ on $X \times X$, $X = \{a, b, c, d\}$.

(b) Let $X = \{2, 3, 4, 5, 6, 7, 8\}$ and define $x\,R\,y$ if $x = y \pmod 3$; that is, if $|x - y|$ is divisible by 3.

2.4. Find all relations on $\{a, b\} \times \{x, y\}$, that are not functional.

2.5. Let S be the divides relation on $\{3, 4, 5\} \times \{4, 5, 6\}$; that is, $x\,S\,y$ if y/x is an integer. List the elements of S and S^{-1}.

2.6. Let A be a set with m elements and B be a set with n elements.

(a) How many relations are there on $A \times B$?

(b) How many functions are there from A to B?

2.7. Define a binary relation D from $\underline{10} = \{1,2,3,4,5,6,7,8,9,10\}$ to $\underline{10}$ as follows: For all x, y, in $\underline{10}$, $x\ D\ y$ if $x < y$ and x divides y. How many edges are there in the directed graph of this relation. Explain.

2.8. This exercise lets you check your understanding of the definitions. In each case below, some information about a function is given to you. Answer the following questions and give reasons for your answers:

- Have you been given enough information to specify the function?
- Can you tell whether or not the function is an injection? a surjection? a bijection?
- If possible, give the function in two-line form.

(a) $f \in \underline{3}^{\{>,<,+,?\}}, \quad f = (3,1,2,3)$.

(b) $f \in \{>,<,+,?\}^{\underline{3}}, \quad f = (?,<,+)$.

(c) $f \in \underline{4}^{\underline{3}}, \quad 2 \to 3, \quad 1 \to 4, \quad 3 \to 2$.

2.9. This exercise lets you check your understanding of the definitions. In each case below, some information about a function is given to you. Answer the following questions and give reasons for your answers:

- Have you been given enough information to specify the function?
- Can you tell whether or not the function is an injection? a surjection? a bijection? If so, what is it?

(a) $f \in \underline{4}^{\underline{5}}, \quad \text{Coimage}(f) = \{\{1,3,5\},\{2,4\}\}$.

(b) $f \in \underline{5}^{\underline{5}}, \quad \text{Coimage}(f) = \{\{1\},\{2\},\{3\},\{4\},\{5\}\}$.

(c) $f \in \underline{4}^{\underline{5}}, \quad f^{-1}(2) = \{1,3,5\}, \quad f^{-1}(4) = \{2,4\}$.

(d) $f \in \underline{4}^{\underline{5}}, \quad \bigl|\text{Image}(f)\bigr| = 4$.

(e) $f \in \underline{4}^{\underline{5}}, \quad \bigl|\text{Image}(f)\bigr| = 5$.

(f) $f \in \underline{4}^{\underline{5}}, \quad \bigl|\text{Coimage}(f)\bigr| = 5$.

2.10. For each of the following definitions, state whether the definition is correct or not correct. If not correct, explain why.

(a) Definition: $f : A \to B$ is one-to-one if $\forall s,t \in A,\ f(s) = f(t)$ implies $s = t$.

(b) Definition: $f : A \to B$ is one-to-one if $\forall s,t \in A,\ s \neq t$ implies $f(s) \neq f(t)$.

(c) Definition: $f : A \to B$ is one-to-one if $\forall s \in A,\ \exists!\, t \in B$ such that $f(s) = t$.

(d) Definition: $f : A \to B$ is one-to-one if $\forall t \in B$, $\exists$ at most one $s \in A$ such that $f(s) = t$.

2.11. Define $g : \mathbb{Z} \to \mathbb{Z}$ by $g(n) = 3n - 1$, where $\mathbb{Z}$ is the set of integers.

(a) Is g one-to-one?

(b) Is g onto?

(c) Suppose that $g : \mathbb{R} \to \mathbb{R}$ and $g(x) = 3x - 1$ for all real numbers x. Is g onto?

2.12. In each case prove or disprove the statement "f is one-to-one."

(a) $f(x) = \frac{x}{x^2+1}$ where $f : \mathbb{R} \to \mathbb{R}$

(b) $f(x) = \frac{2x+1}{x}$ where $f : (\mathbb{R} - \{0\}) \to \mathbb{R}$

(c) $f(x) = \frac{x-1}{x+1}$ where $f : (\mathbb{R} - \{-1\}) \to \mathbb{R}$

2.13. This exercise lets you check your understanding of cycle form. A permutation is given in one-line, two-line or cycle form. Convert it to the other two forms. Give its inverse in all three forms.

(a) (1,5,7,8) (2,3) (4) (6).

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 3 & 7 & 2 & 6 & 4 & 5 & 1 \end{pmatrix}$.

(c) (5,4,3,2,1), which is in one-line form.

(d) (5,4,3,2,1), which is in cycle form. (Assume the domain is $\underline{5}$.)

2.14. Let $S = \{f \mid f : A \to B,\ f \text{ one-to-one}\}$. In each case, find $|S|$.

(a) $|A| = 3$ and $|B| = 3$

(b) $|A| = 3$ and $|B| = 5$

(c) $|A| = m$ and $|B| = n$

2.15. Let $f : X \to Y$, $g : Y \to Z$ and $g \circ f : X \to Z$.

(a) If $g \circ f$ is onto, must f and g be onto?

(b) If $g \circ f$ is one-to-one, must f and g be one-to-one?

2.16. In each case, $f : X \to Y$ is an arbitrary function. Prove or disprove:

(a) For all subsets A and B of X, $f(A \cup B) = f(A) \cup f(B)$.

(b) For all subsets A and B of X, $f(A \cap B) = f(A) \cap f(B)$.

(c) For all subsets A and B of X, $f(A - B) = f(A) - f(B)$.

(d) For all subsets C and D of Y, $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$.

2.17. Let $f : X \to Y$, $g : Y \to Z$ and $g \circ f : X \to Z$. Prove or disprove:

(a) For all subsets $A \subseteq X$, $f^{-1}(f(A)) = A$.

(b) For all subsets $B \subseteq Y$, $f(f^{-1}(B)) = B$.

(c) For all subsets $E \subseteq Z$, $(g \circ f)^{-1}(E)) = f^{-1}(g^{-1}(E))$.

2.18. Let $S = \{f \mid f : A \to B,\ f \text{ onto}\}$. In each case, find $|S|$. Do (a)–(d) without using the general formula in the text.

(a) $|A| = 3$ and $|B| = 2$

(b) $|A| = 3$ and $|B| = 5$

(c) $|A| = 4$ and $|B| = 2$

(d) $|A| = m$ and $|B| = n$

(e) Explain how to use the general formula in the text to solve (d).

2.19. Let $S = \{f \mid f : A \to B,\ |\text{Image}(f)| = k\}$. Suppose that $|A| = m \geq k$ and $|B| = n \geq k$. Prove the formula for $|S|$ given in the text.

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. Which of the following statements is **FALSE**?

(a) $2 \in A \cup B$ implies that if $2 \notin A$ then $2 \in B$.

(b) $\{2, 3\} \subseteq A$ implies that $2 \in A$ and $3 \in A$.

(c) $A \cap B \supseteq \{2, 3\}$ implies that $\{2, 3\} \subseteq A$ and $\{2, 3\} \subseteq B$.

(d) $A - B \supseteq \{3\}$ and $\{2\} \subseteq B$ implies that $\{2, 3\} \subseteq A \cup B$.

(e) $\{2\} \in A$ and $\{3\} \in A$ implies that $\{2, 3\} \subseteq A$.

2. Let $A = \{0, 1\} \times \{0, 1\}$ and $B = \{a, b, c\}$. Suppose A is listed in lexicographic order based on $0 < 1$ and B is in alphabetic order. If $A \times B \times A$ is listed in lexicographic order, then the next element after $((1, 0), c, (1, 1))$ is

(a) $((1, 0), a, (0, 0))$

(b) $((1, 1), c, (0, 0))$

(c) $((1, 1), a, (0, 0))$

(d) $((1, 1), a, (1, 1))$

(e) $((1, 1), b, (1, 1))$

3. Which of the following statements is **TRUE**?

(a) For all sets A, B, and C, $A - (B - C) = (A - B) - C$.

(b) For all sets A, B, and C, $(A - B) \cap (C - B) = (A \cap C) - B$.

(c) For all sets A, B, and C, $(A - B) \cap (C - B) = A - (B \cup C)$.

(d) For all sets A, B, and C, if $A \cap C = B \cap C$ then $A = B$.

(e) For all sets A, B, and C, if $A \cup C = B \cup C$ then $A = B$.

4. Which of the following statements is **FALSE**?

(a) $C - (B \cup A) = (C - B) - A$

(b) $A - (C \cup B) = (A - B) - C$

(c) $B - (A \cup C) = (B - C) - A$

(d) $A - (B \cup C) = (B - C) - A$

(e) $A - (B \cup C) = (A - C) - B$

5. Consider the true theorem, "For all sets A and B, if $A \subseteq B$ then $A \cap B^c = \emptyset$." Which of the following statements is **NOT** equivalent to this statement:

(a) For all sets A^c and B, if $A \subseteq B$ then $A^c \cap B^c = \emptyset$.

(b) For all sets A and B, if $A^c \subseteq B$ then $A^c \cap B^c = \emptyset$.

(c) For all sets A^c and B^c, if $A \subseteq B^c$ then $A \cap B = \emptyset$.

(d) For all sets A^c and B^c, if $A^c \subseteq B^c$ then $A^c \cap B = \emptyset$.

(e) For all sets A and B, if $A^c \supseteq B$ then $A \cap B = \emptyset$.

6. The power set $\mathcal{P}((A \times B) \cup (B \times A))$ has the same number of elements as the power set $\mathcal{P}((A \times B) \cup (A \times B))$ if and only if

(a) $A = B$

(b) $A = \emptyset$ or $B = \emptyset$

(c) $B = \emptyset$ or $A = B$

(d) $A = \emptyset$ or $B = \emptyset$ or $A = B$

(e) $A = \emptyset$ or $B = \emptyset$ or $A \cap B = \emptyset$

7. Let $\sigma = 452631$ be a permutation on $\{1, 2, 3, 4, 5, 6\}$ in one-line notation (based on the usual order on integers). Which of the following is **NOT** a correct cycle notation for σ?

(a) (614)(532)

(b) (461)(352)

(c) (253)(146)

(d) (325)(614)

(e) (614)(253)

8. Let $f : X \to Y$. Consider the statement, "For all subsets C and D of Y, $f^{-1}(C \cap D^c) = f^{-1}(C) \cap [f^{-1}(D)]^c$. This statement is

(a) True and equivalent to:
For all subsets C and D of Y, $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$.

(b) False and equivalent to:
For all subsets C and D of Y, $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$.

(c) True and equivalent to:
For all subsets C and D of Y, $f^{-1}(C - D) = f^{-1}(C) - [f^{-1}(D)]^c$.

(d) False and equivalent to:
For all subsets C and D of Y, $f^{-1}(C - D) = f^{-1}(C) - [f^{-1}(D)]^c$.

(e) True and equivalent to:
For all subsets C and D of Y, $f^{-1}(C - D) = [f^{-1}(C)]^c - f^{-1}(D)$.

9. Define $f(n) = \frac{n}{2} + \frac{1-(-1)^n}{4}$ for all $n \in \mathbb{Z}$. Thus, $f : \mathbb{Z} \to \mathbb{Z}$, $\mathbb{Z}$ the set of all integers. Which is correct?

(a) f is not a function from $\mathbb{Z} \to \mathbb{Z}$ because $\frac{n}{2} \notin \mathbb{Z}$.

(b) f is a function and is onto and one-to-one.

(c) f is a function and is not onto but is one-to-one.

(d) f is a function and is not onto and not one-to-one

(e) f is a function and is onto but not one-to-one.

10. The number of partitions of $\{1, 2, 3, 4, 5\}$ into three blocks is $S(5, 3) = 25$. The total number of functions $f : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4\}$ with $|\text{Image(f)}| = 3$ is

(a) 4×6

(b) 4×25

(c) 25×6

(d) $4 \times 25 \times 6$

(e) $3 \times 25 \times 6$

11. Let $f : X \to Y$ and $g : Y \to Z$. Let $h = g \circ f : X \to Z$. Suppose g is one-to-one and onto. Which of the following is **FALSE**?

(a) If f is one-to-one then h is one-to-one and onto.

(b) If f is not onto then h is not onto.

(c) If f is not one-to-one then h is not one-to-one.

(d) If f is one-to-one then h is one-to-one.

(e) If f is onto then h is onto.

12. Which of the following statements is **FALSE**?

(a) $\{2, 3, 4\} \subseteq A$ implies that $2 \in A$ and $\{3, 4\} \subseteq A$.

(b) $\{2, 3, 4\} \in A$ and $\{2, 3\} \in B$ implies that $\{4\} \subseteq A - B$.

(c) $A \cap B \supseteq \{2, 3, 4\}$ implies that $\{2, 3, 4\} \subseteq A$ and $\{2, 3, 4\} \subseteq B$.

(d) $A - B \supseteq \{3, 4\}$ and $\{1, 2\} \subseteq B$ implies that $\{1, 2, 3, 4\} \subseteq A \cup B$.

(e) $\{2, 3\} \subseteq A \cup B$ implies that if $\{2, 3\} \cap A = \emptyset$ then $\{2, 3\} \subseteq B$.

13. Let $A = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$ and $B = \{a, b, c\} \times \{a, b, c\} \times \{a, b, c\}$. Suppose A is listed in lexicographic order based on $0 < 1$ and B is listed in lexicographic order based on $a < b < c$. If $A \times B \times A$ is listed in lexicographic order, then the next element after $((0, 1, 1), (c, c, c), (1, 1, 1))$ is

(a) $((1, 0, 1), (a, a, b), (0, 0, 0))$

(b) $((1, 0, 0), (b, a, a), (0, 0, 0))$

(c) $((1, 0, 0), (a, a, a), (0, 0, 1))$

(d) $((1, 0, 0), (a, a, a), (1, 0, 0))$

(e) $((1, 0, 0), (a, a, a), (0, 0, 0))$

14. Consider the true theorem, "For all sets A, B, and C if $A \subseteq B \subseteq C$ then $C^c \subseteq B^c \subseteq A^c$." Which of the following statements is **NOT** equivalent to this statement:

(a) For all sets A^c, B^c, and C^c, if $A^c \subseteq B^c \subseteq C^c$ then $C \subseteq B \subseteq A$.

(b) For all sets A^c, B, and C^c, if $A^c \subseteq B \subseteq C^c$ then $C \subseteq B^c \subseteq A$.

(c) For all sets A, B, and C^c, if $A^c \subseteq B \subseteq C$ then $C^c \subseteq B^c \subseteq A$.

(d) For all sets A^c, B, and C^c, if $A^c \subseteq B^c \subseteq C$ then $C^c \subseteq B^c \subseteq A$.

(e) For all sets A^c, B^c, and C^c, if $A^c \subseteq B^c \subseteq C$ then $C^c \subseteq B \subseteq A$.

15. Let $\mathcal{P}(A)$ denote the power set of A. If $\mathcal{P}(A) \subseteq B$ then

(a) $2^{|A|} \leq |B|$

(b) $2^{|A|} \geq |B|$

(c) $2|A| < |B|$

(d) $|A| + 2 \leq |B|$

(e) $2^{|A|} \geq 2^{|B|}$

16. Let $f : \{1,2,3,4,5,6,7,8,9\} \to \{a,b,c,d,e\}$. In one-line notation, $f = (e,a,b,b,a,c,c,a,c)$ (use number order on the domain). Which is correct?

(a) Image$(f) = \{a,b,c,d,e\}$, Coimage$(f) = \{\{6,7,9\},\{2,5,8\},\{3,4\},\{1\}\}$

(b) Image$(f) = \{a,b,c,e\}$, Coimage$(f) = \{\{6,7,9\},\{2,5,8\},\{3,4\}\}$

(c) Image$(f) = \{a,b,c,e\}$, Coimage$(f) = \{\{6,7,9\},\{2,5,8\},\{3,4\},\{1\}\}$

(d) Image$(f) = \{a,b,c,e\}$, Coimage$(f) = \{\{6,7,9,2,5,8\},\{3,4\},\{1\}\}$

(e) Image$(f) = \{a,b,c,d,e\}$, Coimage$(f) = \{\{1\},\{3,4\},\{2,5,8\},\{6,7,9\}\}$

17. Let $\Sigma = \{x, y\}$ be an alphabet. The strings of length seven over Σ are listed in dictionary (lex) order. What is the first string after $xxxxyxx$ that is a palindrome (same read forwards and backwards)?

(a) $xxxxyxy$ (b) $xxxyxxx$ (c) $xxyxyxx$ (d) $xxyyyxx$ (e) $xyxxxyx$

18. Let $\sigma = 681235947$ and $\tau = 627184593$ be permutations on $\{1,2,3,4,5,6,7,8,9\}$ in one-line notation (based on the usual order on integers). Which of the following is a correct cycle notation for $\tau \circ \sigma$?

(a) (124957368)

(b) (142597368)

(c) (142953768)

(d) (142957368)

(e) (142957386)

Answers: **1** (e), **2** (c), **3** (b), **4** (d), **5** (a), **6** (d), **7** (b), **8** (a), **9** (e), **10** (d), **11** (a), **12** (b), **13** (e), **14** (d), **15** (a), **16** (c), **17** (b), **18** (d).