Chapter 7

Relations and Functions

7.1 Ordered pairs

The definition of a set explicitly disregards the order of the set elements, all that matters is who's in, not who's in first. However, sometimes the order is important. This leads to the notion of an *ordered pair* of two elements x and y, denoted (x, y). The crucial property is:

$$(x, y) = (u, v) \text{ if and only if } x = u \text{ and } y = v.$$

This notion can be extended naturally to define an *ordered n-tuple* as the ordered counterpart of a set with n elements.

Give two sets A and B, their *cartesian product* $A \times B$ is the set of all ordered pairs (x, y), such that $x \in A$ and $y \in B$:

$$A \times B = \{(x, y) : x \in A, y \in B\}.$$

Here is a useful special case:

$$A^2 = A \times A = \{(x, y) : x, y \in A\}.$$

And here is a general definition: $A^1 = A$, and for $n \geq 2$,

$$A^n = \big\{(x_1, x_2, \ldots, x_n) : x_1, x_2, \ldots, x_n \in A\big\}.$$

For example, $\mathbb{R}^2$ is the familiar cartesian plane, and $\mathbb{R}^n$ is often referred to as the n-dimensional Euclidean space. If we omit the parentheses and the commas, $\{a, b\}^4$ is comprised of child babble and a 70s pop band:

$$\{aaaa, baba, abab, baaa, baab, aaab, aaba, abaa, abba, bbaa, bbba, bbbb, aabb, abbb, babb, bbab\}.$$

Proposition 7.1.1. $(A \cup B) \times C = (A \times C) \cup (B \times C)$

Proof. Recall that for two sets X and Y, $X = Y$ if and only if $X \subseteq Y$ and $Y \subseteq X$.

Consider any element $(u, v) \in (A \cup B) \times C$. By definition, $u \in A \cup B$ and $v \in C$. Thus, $u \in A$ or $u \in B$. If $u \in A$ then $(u, v) \in A \times C$ and if $u \in B$ then $(u, v) \in B \times C$.

Thus (u, v) is in $A \times C$ or in $B \times C$, and $(u, v) \in (A \times C) \cup (B \times C)$. This proves that $(A \cup B) \times C \subseteq (A \times C) \cup (B \times C)$.

Now consider any element $(u, v) \in (A \times C) \cup (B \times C)$. This implies that $(u, v) \in A \times C$ or $(u, v) \in B \times C$. In the first case $u \in A$ and $v \in C$ and in the second case $u \in B$ and $v \in C$. Thus $u \in A \cup B$ and $v \in C$, which implies $(u, v) \in (A \cup B) \times C$. $\square$

7.2 Relations

Given a set A, a *relation on A* is some property that is either true or false for any ordered pair $(x, y) \in A^2$. For example, "greater than" is a relation on $\mathbb{Z}$, denoted by $>$. It is true for the pair $(3, 2)$, but false for the pairs $(2, 2)$ and $(2, 3)$. In more generality,

Definition 7.2.1. *Given sets A and B, a* relation between A and B *is a subset of $A \times B$.*

By this definition, a relation R is simply a specification of which pairs are related by R, that is, which pairs the relation R is true for. For the relation $>$ on the set $\{1, 2, 3\}$,

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

This notation might look weird because we do not often regard the symbol "$>$" as a meaningful entity in itself. It is, at least from the vantage point of the foundations of mathematics: This symbol is a particular relation.

The common usage of the symbol "$>$" (as in $3 > 2$) is an instance of a useful notational convention: For a relation R, $(a, b) \in R$ can also be specified as aRb. Thus, in the above example, $(2, 1) \in >$ can be written as $2 > 1$. How convenient!

Common mathematical relations that will concern us include $<$, $>$, $\leq$, $\geq$, $=$, $\neq$, $|$, $\equiv_n$, $\subset$, $\subseteq$, etc. For example, the relation $=$ on the set $\mathbb{Z}$ is precisely the set $\{(n, n) : n \in \mathbb{Z}\}$ and the relation $\leq$ on $\mathbb{R}$ is the set $\{(x, x + |y|) : x, y \in \mathbb{R}\}$.

The concept of a relation is as general as the concept of a set, and is not limited to strictly mathematical settings. For instance, we can define the relation *likes* between the set {Anna,Britney,Caitlyn} and the set {Austin,Brian,Carlos}, such that

likes = $\big\{$(Britney,Austin),(Caitlyn,Austin),(Britney, Carlos),(Anna,Austin),(Caitlyn,Brian)$\big\}$.

In this setting we can write *Britney likes Austin*.

7.3 Kinds of relations

A relation R on a set A is called

- *reflexive* if for all $a \in A$, aRa.
- *symmetric* if for all $a, b \in A$, aRb implies bRa.
- *antisymmetric* if for all $a, b \in A$, aRb and bRa implies $a = b$.
- *transitive* if for all $a, b, c \in A$, aRb and bRc implies aRc.

Equivalence relations. A relation that is reflexive, symmetric, and transitive is called an *equivalence relation*. Clearly, the common relation $=$ on the set $\mathbb{R}$, say, is an equivalence relation. Also, we have seen earlier that the congruence relation $\equiv_n$ on the set $\mathbb{Z}$ is reflexive, symmetric, and transitive, thus it is also an equivalence relation. The similarity relation on the set of triangles in the plane is another example.

Equivalence relations are special in that they naturally partition the underlying set into *equivalence classes*. For example, the relation $\equiv_2$ partitions the integers into even and odd ones. These are, respectively, the integers that are related (by $\equiv_2$) to 0, and the ones related to 1. Let's formalize these concepts.

Definition 7.3.1. *A* partition *of a set A is a set $\mathcal{X} \subseteq 2^A \setminus \{\emptyset\}$, such that*

(a) Each $a \in A$ belongs to some $S \in \mathcal{X}$.

(b) If $S, T \in \mathcal{X}$, either $S = T$ or $S \cap T = \emptyset$.

Stated differently, this definition says that the set A is the union of the members of $\mathcal{X}$, and these members are disjoint. Now, given an equivalence relation R on A, the *equivalence class* of $a \in A$ is defined as

$$R[a] = \{b \in A : aRb\}.$$

Theorem 7.3.2. *Let R be an equivalence relation on a set A. Then $\{R[a] : a \in A\}$ is a partition of A.*

Proof. Consider an equivalence relation R on A. Due to reflexivity, every element $a \in A$ belongs to $R[a]$, which implies (a). Now, consider two equivalence classes $R[a]$ and $R[b]$. If aRb, then for any $c \in R[a]$, by transitivity and symmetry, bRc and $c \in R[b]$. This shows $R[a] \subseteq R[b]$. We can symmetrically argue that $R[b] \subseteq R[a]$, which together implies $R[a] = R[b]$.

Otherwise, if $a \not R b$ then consider some $c \in R[a]$. If $c \in R[b]$ then aRc and bRc, which imply, by transitivity and reflexivity, aRb, leading to a contradiction. Thus no element of $R[a]$ belongs to $R[b]$ and $R[a] \cap R[b] = \emptyset$. This shows (b) and concludes the theorem. $\square$

Order relations. A relation that is reflexive, antisymmetric, and transitive is called a *partial order*. The relations $\leq$, $\geq$, and $\mid$ on the set $\mathbb{Z}$, as well as the relation $\subseteq$ on the powerset 2^A of any set A, are familiar partial orders. Note that a pair of elements can be *incomparable* with respect to a partial order. For example, $\mid$ is a partial order on $\mathbb{Z}$, but $2 \nmid 3$ and $3 \nmid 2$. A set A with a partial order on A is called a *partially ordered set*, or, more commonly, a *poset*.

A relation R on a set A is a *total order* if it is a partial order and satisfies the following additional condition:

- For all $a, b \in A$, either aRb or bRa (or both).

For example, the relations $\geq$ and $\leq$ are total orders on $\mathbb{R}$, but $\mid$ is not a total order on $\mathbb{Z}$. Finally, a *strict order* on A is a relation R that satisfies the following two conditions:

- For all $a, b, c \in A$, aRb and bRc implies aRc. (Transitivity.)

- Given $a, b \in A$, exactly one of the following holds (and not the other two): aRb, bRa, $a = b$.

The familiar $<$ and $>$ relations (on $\mathbb{R}$, say) are examples of strict orders.

7.4 Creating relations

There are a few ways to define new relations from existing ones, and we describe two important such ways below.

Restrictions of relations. Here is one notion that is sometimes useful: Given a relation R on a set A, and a subset $S \subseteq A$, we can use R to define a relation on S called the *restriction* of R to S. Denoted by $R|_S$, it is defined as

$$R|_S = \big\{(a, b) \in R : a, b \in S\big\}.$$

Compositions of relations. For three sets A, B, C, consider a relation R between A and B, and a relation S between B and C. The *composition of R and S* is a relation T between A and C, defined as follows: aTc if and only if there exists some $b \in B$, such that aRb and bSc. The composition of R and S is commonly denoted by $R \circ S$.

Note that by this definition, the composition of relations on the same set A is always well-defined. In particular, given a relation R on A we can recursively define $R^1 = R$ and $R^n = R^{n-1} \circ R$ for all $n \geq 2$. Now consider the infinite union

$$T = \bigcup_{i \in \mathbb{N}^+} R^i.$$

This relation T is called the *transitive closure* of R.

Proposition 7.4.1. *Important properties of transitive closure:*

(a) T is transitive.

(b) T is the smallest transitive relation that contains R. (That is, if U is a transitive relation on A and $R \subseteq U$, then $T \subseteq U$.)

(c) If $|A| = n$ then

$$T = \bigcup_{i=1}^{n} R^i.$$

7.5 Functions

The general concept of a function in mathematics is defined very similarly to relations. In fact, as far as the definitions go, functions *are* relations, of a special type:

Definition 7.5.1. *Given two sets A and B, a function $f : A \to B$ is a subset of $A \times B$ such that*

(a) If $x \in A$, there exists $y \in B$ such that $(x, y) \in f$.

(b) If $(x, y) \in f$ and $(x, z) \in f$ then $y = z$.

A function is sometimes called a *map* or *mapping*. The set A in the above definition is the *domain* and B is the *codomain* of f.

A function $f : A \to B$ is effectively a special kind of relation between A and B, which relates every $x \in A$ to *exactly one* element of B. That element is denoted by $f(x)$.

If the above definition is followed rigidly, particular functions should be defined by specifying all the pairs $(x, f(x))$. This is often cumbersome and unnecessary, and we will mostly continue describing a function from A to B as we did before: as a rule for picking an element $f(x) \in B$ for every element $x \in A$. As in, "Consider a function $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = x^2$ for all $x \in \mathbb{R}$."

Kinds of Functions. For a function $f : A \to B$, the set $f(A) = \{f(x) : x \in A\}$ is called the *range* of f. The range is a subset of the codomain but may be different from it. If $f(A) = B$ then we say that f is *onto*. More precisely, a function $f : A \to B$ is a *surjection* (or *surjective*), or *onto* if each element of B is of the form $f(x)$ for at least one $x \in A$.

Today is the day of weird names, so: A function $f : A \to B$ is an *injection* (or *injective*), or *one-to-one* if for all $x, y \in A$, $f(x) = f(y)$ implies $x = y$. Put differently, $f : A \to B$ is one-to-one if each element of B is of the form $f(x)$ for at most one $x \in A$.

As if this wasn't enough: A function $f : A \to B$ is a *bijection* (or *bijective*), or a *one-to-one correspondence* if it is both one-to-one and onto. Alternatively, $f : A \to B$ is a bijection if each element of B is of the form $f(x)$ for exactly one $x \in A$.

Compositions and Inverse Functions. Given two functions $f : A \to B$ and $g : B \to C$ we can define a new function $g \circ f : A \to C$ by $(g \circ f)(x) = g(f(x))$ for all $x \in A$.

One useful function that can be defined for any set A is the *identity function* $i_A : A \to A$, defined by $i_A(x) = x$ for all $x \in A$. We can use identity functions to define *inverse functions*. Specifically, if $f : A \to B$ is a bijection, then its inverse $f^{-1} : B \to A$ is defined so that $f^{-1} \circ f = i_A$ and $f \circ f^{-1} = i_B$. Of course, we haven't shown that f^{-1} even exists or that it is unique, but these properties do hold, assuming that $f : A \to B$ is a bijection. (This assumption is necessary.)

Another result that is sometimes used is the following: If $f : A \to B$ and $g : B \to C$ are bijections then $g \circ f : A \to C$ is a bijection, and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

We omit the proof.

Bijections and cardinality. Bijections allow us to rigorously define when two sets are of the same cardinality:

Definition 7.5.2. *Two sets A and B have the same number of elements if and only if there exists a bijection $f : A \to B$.*

Chapter 10

Binomial Coefficients

10.1 Basic properties

Recall that $\binom{n}{k}$ is the number of k-element subsets of an n-element set, and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{\prod_{i=0}^{k-1}(n-i)}{k!}.$$

The quantities $\binom{n}{k}$ are called *binomial coefficients* because of their role in the *Binomial Theorem*, apparently known to the 11th century Persian scholar Omar Khayyam. Before we state and prove the theorem let us consider some important identities that involve binomial coefficients. One that follows immediately from the algebraic definition is

$$\binom{n}{k} = \binom{n}{n-k}.$$

This also has a nice combinatorial interpretation: Choosing a k-element subset B from an n-element set uniquely identifies the complement $A \setminus B$ of B in A, which is an $(n-k)$-subset of A. This defines a bijection between k-element and $(n-k)$-element subsets of A, which implies the identity.

Another relation between binomial coefficients is called *Pascal's rule*, although it was known centuries before Pascal's time in the Middle East and India:

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

This can be easily proved algebraically:

$$\begin{aligned}
\binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(k-1)!(n+1-k)!} + \frac{n!}{k!(n-k)!} \\
&= \frac{n!k}{k!(n+1-k)!} + \frac{n!(n+1-k)}{k!(n+1-k)!} \\
&= \frac{n!k + n!(n+1-k)}{k!(n+1-k)!} \\
&= \frac{(n+1)!}{k!(n+1-k)!} \\
&= \binom{n+1}{k}.
\end{aligned}$$

Pascal's rule also has a combinatorial interpretation: $\binom{n+1}{k}$ is the number of k-element subsets of an n-element set A. Fix an element $a \in A$. A subset of A either contains a or it doesn't. k-element subsets of A that do not contain a are in fact k-element subsets of $A \setminus \{a\}$ and their number is $\binom{n}{k}$. k-element subsets of A that do contain a bijectively correspond to $(k-1)$-element subsets of $A \setminus \{a\}$, the number of which is $\binom{n}{k-1}$. The identity follows.

Another illuminating identity is the *Vandermonde convolution*:

$$\binom{m+n}{l} = \sum_{k=0}^{l} \binom{m}{k}\binom{n}{l-k}.$$

We only give a combinatorial argument for this one. We are counting the number of ways to choose an l-element subset of an $(m+n)$-element set A. Fix an m-element subset $B \subseteq A$. Any l-element subset S of A has k elements from B and $l-k$ elements from $A \setminus B$, for some $0 \leq k \leq l$. For a particular value of k, the number of k-element subsets of B that can be part of S is $\binom{m}{k}$ and the number of $(l-k)$-element subsets of $A \setminus B$ is $\binom{n}{l-k}$. We can now use the sum principle to sum over the possible values of k and obtain the identity. An interesting special case is

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

It follows from the Vandermonde convolution by taking $l = m = n$ and remembering that $\binom{n}{k} = \binom{n}{n-k}$.

10.2 Binomial theorem

Theorem 10.2.1. *For $n \in \mathbb{N}$ and $x, y \in \mathbb{R}$,*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Proof. By induction on n. When $n = 0$ both sides evaluate to 1. Assume the claim holds for $n = m$ and consider the case $n = m + 1$.

$$\begin{aligned}
(x+y)^{m+1} &= (x+y)\cdot(x+y)^m && (10.1)\\
&= (x+y)\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} && (10.2)\\
&= x\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} + y\cdot\sum_{k=0}^{m}\binom{m}{k}x^k y^{m-k} && (10.3)\\
&= \sum_{k=0}^{m}\binom{m}{k}x^{k+1} y^{m-k} + \sum_{k=0}^{m}\binom{m}{k}x^k y^{m+1-k} && (10.4)\\
&= \sum_{k=1}^{m+1}\binom{m}{k-1}x^{k} y^{m+1-k} + \sum_{k=0}^{m}\binom{m}{k}x^k y^{m+1-k} && (10.5)\\
&= \left(x^{m+1} + \sum_{k=1}^{m}\binom{m}{k-1}x^{k} y^{m+1-k}\right) + \left(y^{m+1} + \sum_{k=1}^{m}\binom{m}{k}x^k y^{m+1-k}\right) && (10.6)\\
&= x^{m+1} + y^{m+1} + \sum_{k=1}^{m}\left(\binom{m}{k-1} + \binom{m}{k}\right)x^k y^{m+1-k} && (10.7)\\
&= x^{m+1} + y^{m+1} + \sum_{k=1}^{m}\binom{m+1}{k}x^k y^{m+1-k} && (10.8)\\
&= \sum_{k=0}^{m+1}\binom{m+1}{k}x^k y^{m+1-k}. && (10.9)
\end{aligned}$$

Here (5) follows from (4) by noting that

$$\sum_{k=0}^{m} f(k) = \sum_{k=1}^{m+1} f(k-1)$$

and (8) follows from (7) by Pascal's rule. The other steps are simple algebraic manipulation. This completes the proof by induction. □

The binomial theorem can be used to immediately derive an identity we have seen before: By substituting $x = y = 1$ into the theorem we get

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n.$$

Here is another interesting calculation: Putting $x = -1$ and $y = 1$ yields

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0.$$

This implies

$$\sum_{k \text{ odd}} \binom{n}{k} = \sum_{k \text{ even}} \binom{n}{k} = 2^{n-1}.$$

This means that the number of odd-size subsets of an n-element set A is the same as the number of even-size subsets, and equals 2^{n-1}. This can be proved by a combinatorial argument as follows: Fix an element $a \in A$ and note that the number of subsets of $A \setminus \{a\}$ is 2^{n-1}. There is a bijective map between subsets of $A \setminus \{a\}$ and odd-size subsets of A, as follows: Map an odd-sized subset of $A \setminus \{a\}$ to itself, and map an even-sized subset $B \subseteq A \setminus \{a\}$ to $B \cup \{a\}$. Observe that this is a bijection and conclude that the number of odd-sized subsets of A is 2^{n-1}. Even-size subsets can be treated similarly, or by noting that their number is 2^n minus the number of odd-size ones.