Functions

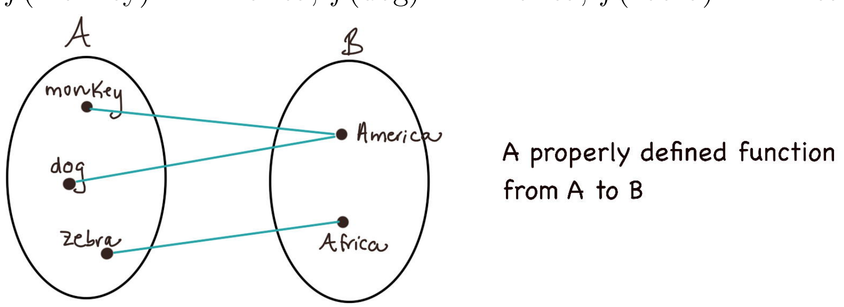
1 First principles

It is quite likely that at some point in your eduction, you have seen functions defined as $f(x) = x^2 + 1$. This is an an example of a function defined over the *real numbers*. Such a function $f(x)$ takes as *input* a real number, x, and *outputs* a real number, often referred to as the y-value. Functions however, can be defined over *any* set, not necessarily just the real numbers. In general, a **function** is simply a mapping (or transformation) of objects from one set to another set.

Definition. *Let A and B be nonempty sets. A function from A to B assigns each value $a \in A$ to exactly one element $b \in B$. It is denoted $f : A \to B$. We write $f(a) = b$ to denote the unique element in B that is assigned by applying f to a.*

The sets can be anything, as long as they are not empty. We could take the set $A = \{monkey, dog, zebra\}$ and $B = \{America, Africa\}$. Then a function would simply have to assign each element in A to an element in B. For example we could set:

$$f(\text{monkey}) = \text{America},\ f(\text{dog}) = \text{America},\ f(\text{zebra}) = \text{Africa}$$

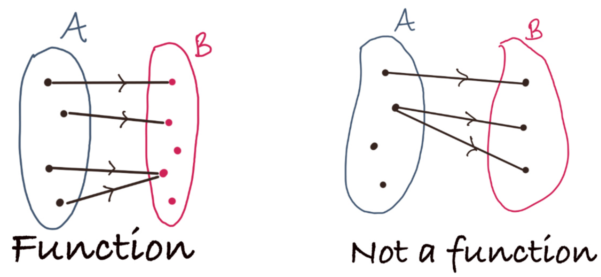


In the above example, we explicitly described the function by stating each assignment. Note that allowing several items to map to the same element in B is allowed ($America$). However *each* element in A has to have an assignment. Thus simply defining

$$f(\text{monkey}) = \text{America},\ f(\text{dog}) = \text{America}$$

is not sufficient to describe the function since there is no assignment for the element $zebra \in A$.

On the other hand, elements of B could be 'left out' in the sense that not every element in B has to have an element in A that is mapped to it. For example if $B = \{America, Africa, Europe, Asia\}$ then the original definition for $f : A \to B$ is still a valid function: $f(\text{monkey}) = \text{America}$, $f(\text{dog}) = \text{America}$, $f(\text{zebra}) = \text{Africa}$.

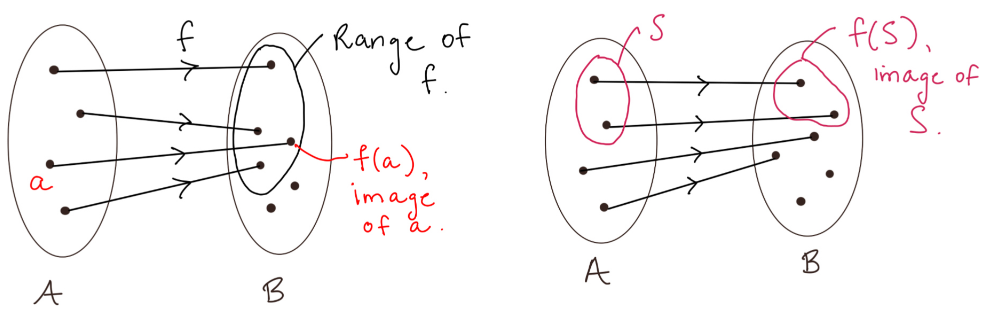


Definition. *If $f : A \to B$, we say that A is the domain of f. If the element a maps to b, i.e. if $f(a) = b$ then b is called the image of a. The range of f is the set of all the images of elements from A.*

For example, if $g : \mathbb{Z} \to \mathbb{Z}$ where $g(x) = x^2$ then the domain is $\mathbb{Z}$, the integers, but the *range* is the subset $\{0, 1, 2, 4, 9, 16, ...\}$ of the integers.

Note that in the definition of a function, the mapping needs to apply to *each* element in the domain, A. Suppose we try to define a function $h : \mathbb{R} \to \mathbb{R}$ where $h(x) = 1/x$. This is *not* a valid function since it cannot be defined for $0 \in \mathbb{R}$. Thus to define this function properly, it's domain must be the set $\mathbb{R} - \{0\}$, which are the real numbers except 0.

Definition. *If $f : A \to B$ and let S be a subset of A. Then we define the image of S under f to be the subset of B that consists of all of the images of the elements from S. The notation is for this set is $f(S)$.*



Suppose for example that $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = 2x+1$, and let $S = \{1, 3, 5\}$. Then $f(S) = \{3, 7, 11\}$.

Example. *Let f be a function from A to B. Let S and T be subsets of A. Show that*

$$f(S \cap T) \subseteq f(S) \cap f(T)$$

Proof: From our section on set theory, we saw how to show that one set is a subset of another. In this particular case, we need to show that any y in the set $f(S \cap T)$ is *also* in the set $f(S) \cap f(T)$. Suppose $y \in f(S \cap T)$. Then y is the *image* of some element $x \in S \cap T$, where $f(x) = y$. Since $x \in S \cap T$, then $x \in S$ and $x \in T$. But the definition of images, this means that $f(x) \in f(S)$ and $f(x) \in f(T)$. Thus $f(x) = y \in f(S) \cap f(T)$.

1.1 Functions as subsets of the Cartesian Product

The function $f : A \to B$ can be also defined in terms of a subset of $A \times B$. For each $a \in A$ we define exactly one pair, (a, b) in $A \times B$. This pair is uniquely defined by $f(a) = b$. In other words, we simply include the pair (a, b) in the subset of $A \times B$ when the element a maps to the element b by the function f. You may have seen functions written like this in the past, often referring the the pair (x, y) as the x and y values of the function. For example, if $f(x) = x^2$ over the integers, then the *pairs* in the subset of $\mathbb{Z} \times \mathbb{Z}$ would include $(1, 1), (-1, 1), (2, 4), (-2, 4), etc.$. In fact, we could refer to this function by saying it is the subset $C = \{(x, x^2) | x \in \mathbb{Z}\}$.

Example 1. *Consider sets $A = B = \mathbb{Z}$ and let $C = \{(x, 0)|x \in \mathbb{Z}\}$. Does this set C define a function?*

Solution: The question we are asked is whether or not the pairs $(x, 0)$ represent possible pairs of a function from $\mathbb{Z}$ to $\mathbb{Z}$. If this were the case, then the function would be defined as $f(x) = 0$, the function that maps all integers to 0. This is indeed a function, although rather simple.

Example 2. *Consider sets $A = B = \{x \in \mathbb{R}| -3 \leq x \leq 3\}$ and let $C = \{(x, y) \in A \times B|x^2 + y^2 = 18\}$ be a subset of $A \times B$. Is this set a function?*

Solution: Is it possible that the pairs (x, y) of C represent the pairs of a function f from A to B? If so, they need to satisfy the definition of a proper function. However, note that the pair $(3, 3) \in C$ and the pair $(3, -3) \in C$, which means that under such a function, $f(3) = 3$ and $f(3) = -3$. This contradicts the definition of a function, namely, that each value $a \in A$ is mapped to *exactly one* value $b \in B$.

Next let's look at some special cases where the function never maps two different values from A to the same element in B.

1.2 Special functions

Definition. *A function $f : A \to B$ is injective (or one-to-one) if $f(a) = f(b)$ implies that $a = b$.*

The natural way to express this is, is to simply say that two different elements in A will always map to two different elements in B. Or *no two* elements in A map to the same element in B. Our example above, $f(x) = x^2$ over the integers, is *not* injective, since $f(1) = f(-1) = 1$.

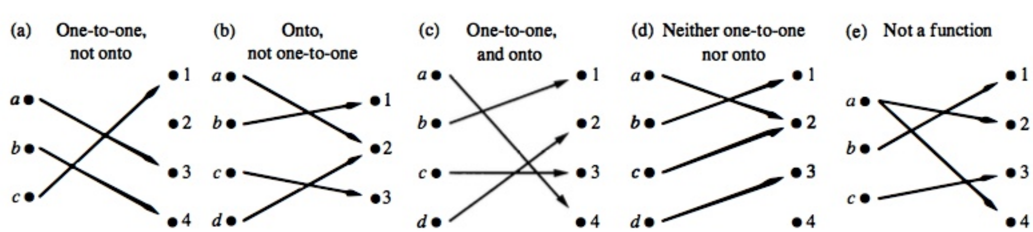
Definition. *A function $f : A \to B$ is surjective (or onto) if and only if for every element $b \in B$ there is an element $a \in A$ where $f(a) = b$.*

Simply put, this means that there are no 'left-overs' in the set B, in the sense that the range of the function is *all* of B. Clearly the function $f(x) = x^2$ is *not* surjective over the integers, since the range is $\{0, 1, 2, 4, 9, 16, ...\}$ which is not all the integers.

Definition. *A bijection is injective and surjective*

The function $f : \mathbb{Z} \to \mathbb{Z}$, $f(x) = x + 1$ is bijective. No two x values map to the same result (injective) *and* any integer y has a pre-image, for example if $y = 5$ then we can set $x = 4$ and $f(4) = 5$.

Here are some examples taken from the Rosen text:



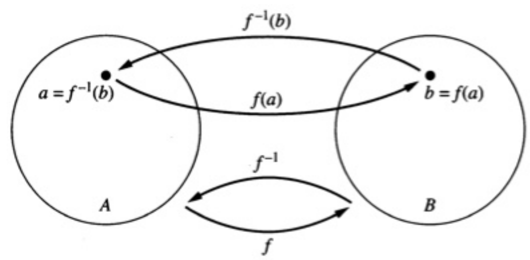
Suppose we wanted to show that a function was *not* injective. In this case, we simply need to find one "counterexample", in other words, a case where $a_1 \neq a_2$ and yet $f(a_1) = b, f(a_2) = b$. From the figure above, we can see that the second function has $f(a) = f(d) = 2$, which is enough to prove that it is *not* injective.

Example 3. *Let $f(x) = -x^3 + 3x + 2$ where $f : \mathbb{R} \to \mathbb{R}$. Is this function injective?*

Solution:. If we can find a value b such that more than one value from A maps to b, then we have shown that the function is not injective. Let's try $b = 2$, and so $-x^3 + 3x + 2 = 2$, or $-x^3 + 3x = 0$ which has several solutions! In fact by factoring we can see that $-x(x^2 - 3) = 0$ which has solutions $x = 0, x = \sqrt{3}, x = -\sqrt{3}$. So certainly this function is *not* injective, since $f(0) = f(-\sqrt{3}) = f(\sqrt{3}) = 2$.

Definition. *If $f : A \to B$ is an injective function, then it is said to be invertible and it's inverse function is $f^{-1}(x) : B \to A$ where $f^{-1}(b) = a$ iff $f(a) = b$.*

Below is a figure taken from Rosen demonstrating how the inverse function maps the element $b \in B$ back to the *preimage* $a \in A$. Clearly if the original function f was *not* injective, there would *not* be a unique element $a \in A$ belonging to $b \in B$, and thus no inverse function could be defined.

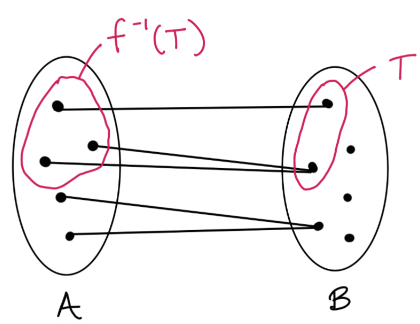


Example 4. *Let $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = 2x$. What is $f^{-1}(8)$?*

Solution: The definition of $f^{-1}(8)$ is simply the value a such that $f(a) = 8$. If the inverse function exists, then this value a is *unique*. In this case, we can see that $a = 4$, since $f(4) = 8$.

We can also define a preimage for an entire *set* $T \subseteq B$.

Definition. *The inverse image of set T under f is the subset $f^{-1}(T)$ of A of all the elements x such that $f(x) \in T$. Note that the inverse image exists for T regardless of whether or not the function is invertible or not!*

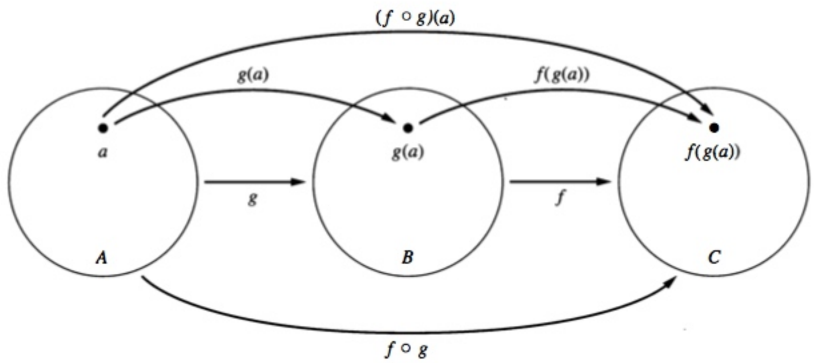


Suppose $f(x) = x^2$ over the real numbers. Then $f(1) = f(-1) = 1$, and thus the function is *not injective* and has no *inverse function*. However, if we define the set $T = \{1, 4, 16\}$, then the *inverse image* of T is $f^{-1}(T) = \{-1, 1, -2, 2, -4, , 4\}$.

Finally we finish this section with the definition of composition:

Definition. *If $g : A \to B$ and $f : B \to C$ then the composite function $f \circ g$ is the function from A to C defined by*

$$f \circ g(x) = f(g(x)) \forall x \in A$$



For example, if $f(x) = x^2$ and $g(x) = x + 1$ over the real numbers, then $g \circ f(x) = x^2 + 1$, whereas $f \circ g(x) = (x+1)^2$. There are cases where two functions may actually give you the same composite function in either order, as in $f(x) = x + 1, g(x) = x - 1$ where $f \circ g(x) = x$ and $g \circ f(x) = x$.

1.3 Floor, ceiling, indicator functions

In many applications of discrete mathematics, such as algorithm analysis in computer science, one often uses the *floor* and *ceiling* functions for counting.

Definition. *The floor function assigns to the real number x the largest integer that is less than or equal to x. The floor function is denoted $\lfloor x \rfloor$. In effect, it "rounds down". The ceiling function assigns to the real number x the smallest integer that is greater than or equal to x. In effect, it "rounds up". The ceiling function is denoted $\lceil x \rceil$.*

For example, $\lfloor 3.4 \rfloor = 3$, $\lceil 3.4 \rceil = 4$, and $\lfloor 3 \rfloor = \lceil 3 \rceil = 3$.

Note that difference between x and its floor/ceiling is not more than one. In fact, if $\lfloor x \rfloor = m$ then $m \le x < m + 1$ and if $\lceil x \rceil = n$ then $n - 1 < x \le n$.

Floor and ceiling functions are often used in counting objects which occur in "wholes". For example, if you can fit 4 people in each train car, and there are 57 people trying to get on the train, then you will need $\lceil \frac{57}{4} \rceil = 15$ train cars.

Let's look at some important properties of floors and ceilings.

Example. *Is it true that $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$ where x, y are real numbers?*

Solution: The above statement is false, since if $x = 1.5$ and $y = 1.5$ then the left side yields 3 and the right side yields 2.

The next theorem tells us that we can add an integer n to x before *or* after we apply the floor function, and the result is the same.

Theorem 1. *For all real numbers x and all integers n,*

$$\lfloor x + n \rfloor = \lfloor x \rfloor + n$$

Proof: Suppose that $\lfloor x \rfloor = m$. Then $m \le x < m + 1$, and if we add n to each part of this inequality we get $m + n \le x + n < m + n + 1$. Since $x + n$ is between $m + n$ and $m + n + 1$, then when we "round it down" it must have floor value $\lfloor x + n \rfloor = m + n$.

Example. *Is it true that $\lceil xy \rceil = \lceil x \rceil \lceil y \rceil$ for real numbers x, y? Is it true that $\lceil xy \rceil \le \lceil x \rceil \lceil y \rceil$ for $x, y \ge 0$?*

Solution: If we take $x = 0.5, y = 1.5$ then the left side above yiels $\lceil 0.75 \rceil = 1$ and the right side yields 2. So we have shown that the first equality is not true in general. It looks like the second inequality is true. To prove it, suppose that $\lceil x \rceil = m$ and $\lceil y \rceil = n$. So that means $\lceil x \rceil \lceil y \rceil = mn$. We also know by definition of the ceiling function that $x \leq m$ and $y \leq n$. So $xy \leq mn$, and thus $\lceil xy \rceil \leq mn$. Putting these two facts together yields $\lceil xy \rceil \leq \lceil x \rceil \lceil y \rceil$.

A similar result exists for the floor function.

The last function we look at in this section is the $indicator function$ defined for a set A.

Definition. *Suppose that A is some set, which is a subset of the universal set. Define the indicator function f_A of the set A as the function that maps each element $x \in A$ to 1: in other words, $f_A(x) = 1$ for all $x \in A$ and $f_A(x) = 0$ for all $x \notin A$.*

For example, if $A = \{1, 2, 3\}$ over the integers, then $f_A(1) = f_A(2) = f_A(3) = 1$, and all other x values yield 0. Suppose that A, B are both sets. Consider the indicator function on their intersection. It should be fairly obvious that $f_{A\cap B}(x) = f_A(x)f_B(x)$. What can you say about $f_{A\cup B}(x)$?

2 Cardinality

The concept of cardinality is very important in computer science - the basic construction of many algorithms often relies on a correct notion of the cardinality of sets. Recall that in the section on set theory, we defined cardinality, $|S|$, for *finite* sets to be the *number of elements* in the set S. For sets of *infinite* size, we *compare* the cardinality of two sets A and B, by creating mappings between their elements. This gives us an idea of the *comparative* size of infinite sets. We do *not* simply say $|A| = \infty$ and $|B| = \infty$ and thus they have the "same" number of elements. In fact, we shall cases where two infinite sets have the same cardinality even though it may seem to us that one set is "larger" than the other.

We start out by discussing the set of *natural numbers*, $\mathbb{N} = \{1, 2, 3, \ldots\}$. This set is clearly infinite. It is also a countable set, which means simply we can *count* the elements as $1, 2, 3...$ etc.

Other infinite sets can also be countable, as long as we can *count* the elements in that set. Counting the elements is like enumerating the elements as the *first,second, third..* etc. In order to guarantee that this can be done, we simply need the following requirement:

Definition. *A set S is called countable if it is either finite OR if there exists a bijection from S to $\mathbb{N}$ in which case it is called countably infinite.*

Let's look at some sets that are countable, even though they may at first seem *larger* than the natural numbers.

Example. *The integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ are countable*

Proof: We need to find a way to 'count' these elements, and simply assuming we can count the positive integers and then 'come back' and count the negative integers is not valid - because there are an infinite number of positive integers. Instead we need to order them in a systematic way that ensures they all get counted. One system would be to start with 0, then count -1,1, then count -2,2, etc. Thus the mapping to the natural numbers would be in the order:

$$\mathbb{Z} = \{0, -1, 1, -2, 2, -3, 3, \ldots\}$$

Example. *The positive rational numbers $\mathbb{Q}$ are countable*

Proof: The positive rational numbers consist of all elements of the form p/q where p, q are positive integers, and $q \neq 0$. How can we find a way to systematically count all of these possible fractions? The result is due to Cantor, a Russian mathematician, in 1874. He devised to count the elements in a *diagonal* manner:

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{1}, \frac{2}{2}, \frac{1}{4}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \ldots\ldots$$

where the pattern is based on ordering the fractions by the total *sum* of $p+q$. If we count the positive rationals in this way, we have *not quite* created a valid bijective mapping from $\mathbb{Q}^+$ to $\mathbb{N}$ since the fractions $1/2$ and $2/4$ will appear twice on the list although they are the same rational number. We could simply 'skip' those rational numbers that we have already counted. This would indeed create a bijective function, although we can't write is explicity, but it does show that one exists.

By similar arguments, one can show that $\mathbb{N} \times \mathbb{N}$ is countable.

An example of an **uncountable** set is the *irrational numbers* (we are not going to see the proof here). That means that there is *no* bijective function from $\mathbb{N}$ to the irrationals. And since the real numbers include the irrationals, then the **real numbers are also uncountable**.

We can also directly *compare* the cardinality of two sets in the following way:

Definition. *We say that sets A and B have the same cardinality and write $|A| = |B|$ if there is a bijection from A to B. Similary, we write $|A| \leq |B|$ if there is an injective function from A to B.*

This definition can create some rather unnatural notions of sets size.

Example 5. *Show that the integers and the set of even integers have the same cardinality*

Solution:. This fact may seem quite unnatural. Most of us would probably conclude that the set of even integers, let's call it $S = \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\}$ is "smaller" that the set of integers, $\mathbb{Z} = \{\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$. However, we can define a bijective function $f : \mathbb{Z} \to S$ by $f(x) = 2x$. This mapping confirms that indeed $|\mathbb{Z}| = |S|$.