Mathematics for Computer Scientists

Lecture notes for the module G51MCS

Venanzio Capretta
University of Nottingham
School of Computer Science

Chapter 3

Arithmetic

There are several sets of numbers, from the most simple *natural numbers*, which are used to count objects, to the *real numbers*, which represent points on a straight line.

We consider four number systems, each contained in the next one, and indicate each of them with a symbol in blackboard font:

- $\mathbb{N}$: the set of *natural numbers*, $0, 1, 2, 3, \ldots$, the counting numbers;
- $\mathbb{Z}$: the set of *integer numbers*, comprising the naturals and their negations, $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$.
- $\mathbb{Q}$: the set of *rational numbers*, those that can be expressed by *fractions*, $1/2, 5/21, -15/4$.
- $\mathbb{R}$: the set of *real numbers*, all those that can be expressed by a decimal notation, even infinite, like $5.129641064\cdots$. Examples of real numbers that are not rational are $\sqrt{2}$ and π. Real numbers can also be described as those that denote the points on a straight line, once we choose a point of origin and a unit of measure.

Some properties of numbers will be the same independently of what set we are considering, but others will be true only in certain number systems.

3.1 Order relations

There are four order relations that are used to compare two numbers: Suppose x and y are two numbers belonging to any of the number systems above. We can compare them in any of the following ways:

- $x \leq y$: *x is smaller than y or equal to y.* We can also express the same relation by saying *x is at most y.*

- $x < y$: *x is (strictly) smaller than y.* This relation could be defined in terms of the previous one by:

$$x < y \Leftrightarrow x \leq y \wedge x \neq y.$$

x is smaller than y is equivalent to *x is at most y and not equal to y.*

- $x \geq y$: *x is larger than y or equal to y.* Equivalently we can say *x is at least y.* Clearly this is the inverse relation to the first one:

$$x \geq y \Leftrightarrow y \leq x.$$

- $x > y$: *x is (strictly) larger than y.* This is the inverse relation to the second one:

$$x > y \Leftrightarrow y < x.$$

These order relations satisfy some basic properties, which are the same for all number systems.

Reflexivity: $x \leq x$. Every number is less than itself or equal to itself.

Antisymmetry: $x \leq y \wedge y \leq x \Rightarrow x = y$. If, of two numbers, the first is less than or equal to the second and, vice versa, the second is less than or equal to the first, then the two numbers must be equal.

Transitivity: $x \leq y \wedge y \leq z \Rightarrow x \leq z$. If, of three numbers, the first is less than or equal to the second and the second is less than or equal to the third, then the first is less than or equal to the third.

Totality: $x \leq y \vee y \leq x$. Of two numbers, one must be less or equal to the other or vice versa.

Some other properties of the order relation are specific to one number system. For example consider the following proposition:

$$x < y \Rightarrow x + 1 \leq y.$$

It states that if x is strictly smaller than y, then $x + 1$ is at most y. This is a true fact if x and y are integers, since two distinct integers must be separated by a distance of at least 1. But it is not true for fractions and real numbers, that may have a smaller distance. Therefore, the proposition is true in $\mathbb{N}$ and $\mathbb{Z}$, but it is false in $\mathbb{Q}$ and $\mathbb{R}$.

3.2 Absolute Value, Floor and Ceiling functions

Each of the four number systems in the previous section is contained in the next one. We use the symbol $\subseteq$ to represent the relation *being contained.* So the previous fact is written symbolically so:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}.$$

We can go the other way and "force" a number in one of the higher systems into a lower one. First of all, we can turn any number into a positive one or zero by taking its *absolute value*, that is, its value *without the sign.* We enclose a number between vertical bars to denote its absolute value, so: $|-5| = 5, |+3| = 3, |0| = 0, |-2.5| = 2.5, |7| = 7$. For a precise mathematical definition, we say that the absolute value of a number x is x itself if it is positive or zero and it is its opposite if it is negative:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

We can approximate a rational or real number with an integer from below, using the *floor* function, or from above, using the *ceiling* function.

The *floor* of a number x is denoted by $\lfloor x \rfloor$ and it is the largest integer that is smaller than or equal to x. For example:

$$\lfloor 1.5 \rfloor = 1, \lfloor -1.5 \rfloor = -2, \lfloor 0 \rfloor = 0, \lfloor 5 \rfloor = 5.$$

We can express the defining property of the floor function by the following formula:

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1.$$

The *ceiling* of a number x is denoted by $\lceil x \rceil$ and it is the smallest integer that is larger than or equal to x. For example:

$$\lceil 1.5 \rceil = 2, \lceil -1.5 \rceil = -1, \lceil 0 \rceil = 0, \lceil 5 \rceil = 5.$$

We can express the defining property of the ceiling function by the following formula:

$$\lceil x \rceil - 1 < x \leq \lceil x \rceil.$$

3.3 Divisibility

From now on we will concentrate on natural numbers. So all the operations and relations we talk about will be in $\mathbb{N}$. Another important relation, beside the *less than or equal* relation, is *divisibility.* We say that a number n *divides* a number m if we can write m as a multiple of n: that is, if there is some other number h such that $m = h \cdot n$. We use the symbolic notation $n \mid m$ to express this fact. So, in formulas:

$$n \mid m \Leftrightarrow m = h \cdot n \text{ for some number } h.$$

For example $3 \mid 18$ because we can write $18 = 6 \cdot 3$, $7 \mid 35$ because we can write $35 = 5 \cdot 7$, but NOT $5 \mid 31$ because it is impossible to find a multiple of 5 that's equal to 31. In this last case we write $5 \nmid 31$ to express the fact that 5 does not divide 31.

The divisibility relation satisfies some of the same properties that are true for the *less than or equal* relation, in particular reflexivity, antisymmetry, and transitivity.

Reflexivity: $n \mid n$. Every number divides itself: In fact, we can write $n = 1 \cdot n$.

Antisymmetry: $n \mid m \wedge m \mid n \Rightarrow n = m$. If, of two numbers, the first divides the second and, vice versa, the second divides the first, then the two numbers must be equal: The first hypothesis tells us that $m = h \cdot n$ for some number h, the second tells us that $n = k \cdot m$ for some number k; but then it must be $n = k \cdot h \cdot n$; this can happen (in $\mathbb{N}$) only if $h = k = 1$ and so $n = k \cdot m = 1 \cdot m = m$. Notice that this line of reasoning is valid only in $\mathbb{N}$: for example, in $\mathbb{Z}$ it would be possible that $h = k = -1$, and therefore antisymmetry is not necessarily true.

Transitivity: $n \mid m \wedge m \mid p \Rightarrow n \mid p$. If, of three numbers, the first divides the second and the second divides the third, then the first divides the third: The first hypothesis tells us that $m = h \cdot n$ and the second that $p = k \cdot m$; therefore $p = (k \cdot h) \cdot n$, so p is a multiple of n and the conclusion is true.

But not all properties of the *less or equal* relation are also true of the *divisibility* relation. In fact, the one we called *totality* is false: given two numbers n and m, it is not necessary that one of them divides the other or vice versa. In fact, if you take $n = 2$ and $m = 3$, you easily see that 2 doesn't divide 3 and 3 doesn't divide 2.

Important. There is a subtle point that we have ignored above: which numbers divide 0 and which are divided by it? First of all, a number n divides 0 if 0 can be written as a multiple of n: but this is very easy, since $0 = 0 \cdot n$. Therefore:

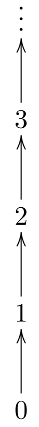
$$n \mid 0 \quad \text{for every number n.}$$

On the other hand, 0 divides a number n only if we can write n as a multiple of 0: $n = h \cdot 0$ for some number h. But obviously this implies that $n = 0$, independently of what value h has. So 0 only divides itself.

The reasoning that we used to show that antisymmetry holds for the divisibility relation only works if both n and m are different from 0. Check its steps carefully and verify that something doesn't work if either n or m happens to be zero. Therefore, if we want to have a rigorous proof of antisymmetry, we should analyse separately the cases when $n = 0$ or $m = 0$. You can verify for yourself that in those cases antisymmetry still holds.

We can use a diagram to give an image of an order relation. Just write the elements/numbers on the page and connect them with arrows: draw an arrow from a number x to another number y if the relation holds between the two.

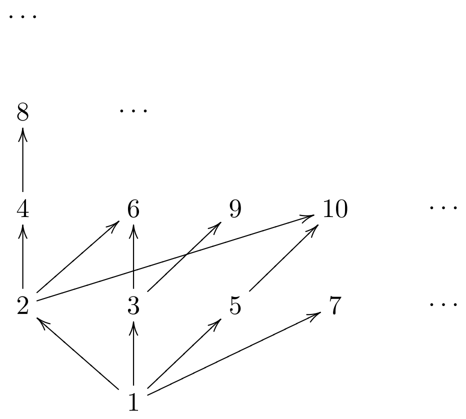
In the case of the order relation *less than or equal* ($\leq$), part of the diagram looks like this:



The number 0 is at the bottom because it is the smallest one. There is an arrow from 0 to 1 because $0 \leq 1$, there is an arrow from 1 to 2 because $1 \leq 2$, and so on. Of course, it is also true that $0 \leq 2$, but we don't draw an arrow between them because it is already implicit from the arrows from 0 to 1 and from 1 to 2 (the transitivity property guarantees this). We also don't draw any arrow from each of the elements to itself to express the fact that $n \leq n$, because that is implicitly guaranteed by the reflexivity property.

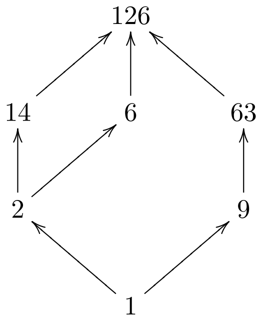
You can see that the relation $\leq$ gives rise to a diagram that can be drawn in a straight ascending line. This happens because of the *totality* property, that ensures that all the elements can be ordered one after the other. This property is therefore sometimes called *linearity.*

As we have seen, the divisibility relation is not total, therefore it generates a much more interesting diagram:



Try to extend this diagram up to the number 20.

If I give you a finite set of numbers, let's say $\{1, 2, 6, 9, 14, 63, 126\}$, you should be able to draw a diagram of the divisibility relation for them, like this:



This kind of drawing is called a *Hasse diagram*.

3.4 Quotient and Remainder

When a number doesn't divide another exactly, we can still compute an approximate result of the division, with some left-over part that could not be divided. For example, if we try to divide 30 by 7 we cannot get an exact result (using only natural numbers), but we can approximate 30 by $4 \cdot 7$ which gives 28, we have 2 left over from the division:

$$30 = 4 \cdot 7 + 2.$$

We say that 4 is the *quotient* of the division and 2 is the *remainder*. In general we are trying to make the remainder as small as possible: we should always ensure that it is smaller than the value we divide by.

Given two natural numbers m and n, with $n \neq 0$, the *quotient* and *remainder* of the division of m by n are the natural numbers q and r such that:

$$m = q \cdot n + r \qquad \text{with } 0 \leq r < n.$$

3.5 Recursive Functions

The set of natural numbers, $\mathbb{N}$, can be enumerated one element after the other, starting with zero and adding one at each step: $0, 1, 2, 3, \ldots$. This obvious fact has very powerful consequences: whenever we want to define a function or an operation on $\mathbb{N}$ we can also do it stepwise, first stating the value of the function on the argument 0 and then prescribing how to obtain the next value from the previous one; similarly, if we want to prove some property of all natural numbers, we can first prove that it holds for zero and then show how to deduce that it holds for the next number if we already know it for the previous one. These procedures are called *recursion* and *induction*: they are among the most powerful tools in computer science and mathematics.

How do we define the exponential function with base 2 on the natural numbers? That is, how is 2^n defined as a function of n? It is customary to say that 2^n is *the product of n copies of* 2 or, symbolically,

$$2^n = \underbrace{2 \times \cdots \times 2}_{n \text{ times}}.$$

The use of three dots to tell the reader to repeat a certain operation a number of times is very common, but it is not a rigorous mathematical notation. Among mathematicians, it is understood that the three-dots notation is a shortened way of hinting at the proper *recursive* definition:

$$\begin{aligned} &2^0 = 1 \\ &2^n = 2 \times 2^{n-1} \qquad \text{if } n > 0. \end{aligned}$$

This gives a recipe to compute the exponential 2^n for every number n: if n is zero, then the result is 1; if n is larger than zero, then first compute the exponential for the previous number, 2^{n-1}, and then multiply that result by 2. Using this recipe we can compute the exponential on every element of $\mathbb{N}$:

$$\begin{array}{ll} 2^0 = 1 & \text{by definition} \\ 2^1 = 2 \times 2^{1-1} = 2 \times 2^0 = 2 \times 1 = 2 & \text{by definition and previous case} \\ 2^2 = 2 \times 2^{2-1} = 2 \times 2^1 = 2 \times 2 = 4 & \text{by definition and previous case} \\ 2^3 = 2 \times 2^{3-1} = 2 \times 2^2 = 2 \times 4 = 8 & \text{by definition and previous case} \\ 2^4 = 2 \times 2^{4-1} = 2 \times 2^3 = 2 \times 8 = 16 & \text{by definition and previous case} \\ \quad \vdots & \end{array}$$

A similar definition is used to compute the *factorial* of a number. The factorial of a natural n is *the product of all the numbers from* 1 *to* n. It is denoted by $n!$. Using the three dots notation:

$$n! = 1 \times \cdots \times n.$$

With a proper recursive definition, this becomes:

$$\begin{aligned} &0! = 1 \\ &n! = n \times (n-1)! \qquad \text{if } n > 0. \end{aligned}$$

Again, we can use this definition as a recipe to compute the factorials of all natural numbers:

$$\begin{array}{ll} 0! = 1 & \text{by definition} \\ 1! = 1 \times (1-1)! = 1 \times 0! = 1 \times 1 = 1 & \text{by definition and previous case} \\ 2! = 2 \times (2-1)! = 2 \times 1! = 2 \times 1 = 2 & \text{by definition and previous case} \\ 3! = 3 \times (3-1)! = 3 \times 2! = 3 \times 2 = 6 & \text{by definition and previous case} \\ 4! = 4 \times (4-1)! = 4 \times 3! = 4 \times 6 = 24 & \text{by definition and previous case} \\ \quad \vdots & \end{array}$$

The pattern is clear. A recursive definition consists of two cases: the first, *the basis of the recursion*, gives the value of the function on zero; the second, *the recursive step*, tells us how to compute the value of the function on a non-zero number from the value of the function on the previous number. The general pattern looks like this (here we write $f(n)$ to indicate the value of the function that we are defining on a number n):

$$\begin{aligned} f(0) &= \textsf{some value (a natural number)} \\ f(n) &= \textsf{some expression containing } f(n-1) \qquad \text{if } n > 0. \end{aligned}$$

Let us try it out by defining our own original function on natural numbers, which we call $\textsf{myFun}$:

$$\begin{aligned} \textsf{myFun}(0) &= 1 \\ \textsf{myFun}(n) &= \textsf{myFun}(n-1) + 2 \times n + 1 \qquad \text{if } n > 0. \end{aligned}$$

We can use the definition of our function to compute it on all natural numbers:

$$\begin{aligned} \textsf{myFun}(0) &= 1 && \text{by definition} \\ \textsf{myFun}(1) &= \textsf{myFun}(1-1) + 2 \times 1 + 1 && \text{by definition and} \\ &= \textsf{myFun}(0) + 3 = 1 + 3 = 4 && \text{previous case} \\ \textsf{myFun}(2) &= \textsf{myFun}(2-1) + 2 \times 2 + 1 && \text{by definition and} \\ &= \textsf{myFun}(1) + 5 = 4 + 5 = 9 && \text{previous case} \\ \textsf{myFun}(3) &= \textsf{myFun}(3-1) + 2 \times 3 + 1 && \text{by definition and} \\ &= \textsf{myFun}(2) + 7 = 9 + 7 = 16 && \text{previous case} \\ \textsf{myFun}(4) &= \textsf{myFun}(4-1) + 2 \times 4 + 1 && \text{by definition and} \\ &= \textsf{myFun}(3) + 9 = 16 + 9 = 25 && \text{previous case} \\ &\vdots \end{aligned}$$

There is a more liberal form of definition by recursion. Instead of using just the previous result, we could allow ourselves to use any or all of the values of the function on smaller numbers. In other words, we could use the following scheme of definition:

$$\begin{aligned} g(0) &= \textsf{some value (a natural number)} \\ g(n) &= \textsf{some expression containing } f(k) \text{ for any } k < n \qquad \text{if } n > 0. \end{aligned}$$

We could also allow ourselves to give more than one starting value. For example, we could give base cases both for zero and for one, like this:

$$\begin{aligned} g(0) &= \textsf{some value (a natural number)} \\ g(1) &= \textsf{some other value} \\ g(n) &= \textsf{some expression containing } f(k) \text{ for any } k < n \qquad \text{if } n > 1. \end{aligned}$$

In this case the recursive step starts being applied when the value of n is at least 2. We can have as many base cases as we wish.

One example of such a recursive definition is the famous *Fibonacci sequence*. Imagine you're a rabbit breeder! You start out with a pair of newborn rabbits.

Adult rabbits reproduce once a month: a pair will produce another pair each month. Newborns do not reproduce for the first month but start giving offspring on the second month after their birth. The mathematical question is: how many rabbits will you have at the nth month? Let's say that the month when you bought your first pair of newborn rabbits is month zero. Let's write $\mathsf{Fib}(n)$ to denote the number of rabbit pairs you have at month n. Then $\mathsf{Fib}(0) = 1$ because initially you have only one pair. But also $\mathsf{Fib}(1) = 1$, because the newborn pair does not reproduce on the first month. On the second month, your pair, now adult, will give a pair of offspring: $\mathsf{Fib}(2) = 2$. On the third month, your original pair will give another pair of offspring, while the newborn pair from month 2 does not yet reproduce: $\mathsf{Fib}(3) = 3$. The next month you have two adult pairs which reproduce, giving two more newborn pairs, while the pair that was born on month 3 does not yet reproduce: $\mathsf{Fib}(4) = 5$. And so on. Looking for a general formula, we notice that at month n all the pairs that were present at month $n - 1$ are still there, plus there will be a batch of newborn pairs equal to the number of pairs present at month $n - 2$ (because those are the ones that are adult). This translates readily to a recursive definition:

$$\begin{aligned}
&\mathsf{Fib}(0) = 1\\
&\mathsf{Fib}(1) = 1\\
&\mathsf{Fib}(n) = \mathsf{Fib}(n-1) + \mathsf{Fib}(n-2) \qquad \text{if } n > 1.
\end{aligned}$$

You can use it to compute how many pairs of rabbits you will have at every month:

$$\begin{aligned}
\mathsf{Fib}(0) &= 1 && \text{by definition}\\
\mathsf{Fib}(1) &= 1 && \text{by definition}\\
\mathsf{Fib}(2) &= \mathsf{Fib}(2-1) + \mathsf{Fib}(2-2) && \text{by definition and}\\
&= \mathsf{Fib}(1) + \mathsf{Fib}(0) = 1 + 1 = 2 && \text{previous cases}\\
\mathsf{Fib}(3) &= \mathsf{Fib}(3-1) + \mathsf{Fib}(3-2) && \text{by definition and}\\
&= \mathsf{Fib}(2) + \mathsf{Fib}(1) = 2 + 1 = 3 && \text{previous cases}\\
\mathsf{Fib}(4) &= \mathsf{Fib}(4-1) + \mathsf{Fib}(4-2) && \text{by definition and}\\
&= \mathsf{Fib}(3) + \mathsf{Fib}(2) = 3 + 2 = 5 && \text{previous cases}\\
\mathsf{Fib}(5) &= \mathsf{Fib}(5-1) + \mathsf{Fib}(5-2) && \text{by definition and}\\
&= \mathsf{Fib}(4) + \mathsf{Fib}(3) = 5 + 3 = 8 && \text{previous cases}\\
\mathsf{Fib}(6) &= \mathsf{Fib}(6-1) + \mathsf{Fib}(6-2) && \text{by definition and}\\
&= \mathsf{Fib}(5) + \mathsf{Fib}(4) = 8 + 5 = 13 && \text{previous cases}\\
&\vdots
\end{aligned}$$

Definition by recursion is so useful that it is the basis of some programming languages, called *functional languages*, for example Haskell. In fact, the definitions we gave above are essentially Haskell programs.

3.6 Euclid's Algorithm

One of the oldest algorithms ever devised is a classic example of recursive definition: Euclid's method to compute the greatest common divisor of two natural numbers. Let's recall some definitions first.

Definition 2 *The* quotient *and* remainder *of the division of a natural number m by another natural number n, with $n > 0$, are the values q and r such that:*

$$m = q \times n + r \qquad 0 \leq r < n.$$

We denote the quotient q by $\mathsf{quot}(m, n)$ *or $m \div n$ and the remainder r by* $\mathsf{rem}(m, n)$.

The standard procedure to compute the quotient and remainder of two large numbers is *long division.*

Definition 3 *We say that a number n is a* divisor *of m if* $\mathsf{rem}(m, n) = 0$*, that is, if m can be written as a multiple of n: $m = q \times n$.*

Definition 4 *The* greatest common divisor *of two numbers m and n, if at least one of the two is different from zero, is the largest natural number that divides both n and m. It is denoted by* $\mathrm{gcd}(m, n)$.

Consider for example the two natural numbers 28 and 70. Let's compute their greatest common divisor. The divisors of 28 are: $1, 2, 4, 7, 14, 28$. The divisors of 70 are: $1, 2, 5, 7, 10, 14, 35, 70$. The largest common divisor is 14, so $\mathrm{gcd}(28, 70) = 14$.

Observation: Can you understand why, in the definition of gcd, we had to put in the requirement that at least one of the two numbers is different from zero? What happens if we try to find $\mathrm{gcd}(m, n)$ when both $m = 0$ and $n = 0$? Does the definition still make sense?

If m and n are very large numbers, then it becomes impractical to compute their gcd by listing all their divisors. There is a much more efficient way to do it: Euclid's algorithm. It relies on the following fact:

Theorem 5 *If m and n are two natural numbers and $n > 0$; then* $\mathrm{gcd}(m, n) = \mathrm{gcd}(n, r)$ *where* $r = \mathsf{rem}(m, n)$.

Proof. In fact, we know, by the very definition of remainder, that $m = q \times n + r$, so $r = m - q \times n$. What we need to prove is that all common divisors of n and m are also divisors of r and, vice versa, that all common divisors of n and r are also divisors of m. This will allow us to replace m with r when computing the gcd.

So, first let's assume that u is a common divisor of m and n, that is, $m = h \times u$ and $n = k \times u$ for some numbers h and k. But then

$$r = m - q \times n = h \times u - q \times k \times u = (h - q \times k) \times u,$$

so u is also a divisor of r.

Vice versa, assume that v is a common divisor of n and r, that is, $n = i \times v$ and $r = j \times v$ for some numbers i and j. But then

$$m = q \times n + r = q \times i \times v + j \times v = (q \times i + j) \times v,$$

so v is also a divisor of m.

We can conclude that the common divisors of n and r are the same as the common divisors of m and n, therefore $\gcd(m, n) = \gcd(n, r)$. □

The previous theorem gives us a way to reduce the size of the numbers of which we want to compute the gcd. In fact, it allows us to write a recursive algorithm to compute it:

Euclid's Algorithm:

$$\begin{array}{ll} \gcd(m, 0) = m & \text{for any number } m > 0 \\ \gcd(m, n) = \gcd(n, r) \text{ where } r = \mathsf{rem}(m, n) & \text{if } n > 0, \text{ for any number } m. \end{array}$$

Contrary to our previous examples, gcd is a function of two arguments, m and n. But this is not a difficulty: the first argument m can be any number. It is the second one, n, that allows us to do recursion. Remember that the remainder r has the property that $0 \leq r < n$, so r is always smaller than n. Therefore, by changing $\gcd(m, n)$ to $\gcd(n, r)$ we have decreased the second argument, and we are assured that the computation will eventually come to a stop.

Here is an example of a run of Euclid's algorithm. We are going to compute $\gcd(1455, 630)$. On the computation below, I have written on the right-hand side the computations of quotient and remainder to clarify each step:

$$\begin{array}{rlr} \gcd(1455, 630) & = \gcd(630, 195) & 1455 = 2 \times 630 + 195 \\ & = \gcd(195, 45) & 630 = 3 \times 195 + 45 \\ & = \gcd(45, 15) & 195 = 4 \times 45 + 15 \\ & = \gcd(15, 0) & 45 = 3 \times 15 + 0 \\ & = 15. & \end{array}$$

3.7 Summations

One kind of recursive definition is so important that it has its own special notation: summation. Suppose we want to add up a sequence of values, each given by a particular case of an expression. A simple case is the addition of all natural numbers up to a given value:

$$0 + 1 + 2 + 3 + \cdots + n.$$

For example, if $n = 7$ we get: $0 + 1 + 2 + 3 + 4 + 5 + 6 + 7 = 28$. As we observed earlier, the three dots notation is a sign that we are doing recursion. Let's call $S(n)$ the result of this summation. We can then define it as a recursive function:

$$\begin{array}{ll} S(0) = 0 & \\ S(n) = S(n - 1) + n & \text{if } n > 0. \end{array}$$

There is a standard notation to represent this function and functions similar to it. We can denote $S(n)$ by the *sigma notation:*

$$\sum_{i=0}^{n} i.$$

The symbol Σ is the capital Greek letter sigma. It tells us that we are computing the addition of several values. The argument of the sigma is i and it is an expression giving a generic element of the sequence that we are adding up. Under the sigma we write $i = 0$ to indicate that the first element of the addition is computed at the value zero for i. Above the sigma we write n to indicate that the last element of the addition is computed at the value n for i. For example, when $n = 7$, we have:

$$\sum_{i=0}^{7} i = 0 + 1 + 2 + 3 + 4 + 5 + 6 + 7 = 28.$$

Do you remember the first original recursive function that we defined, myFun? Look at it again and you will realize that it is also a summation: we start with the basis $\mathsf{myFun}(0) = 1$ and then, at each successive step, we add $2n + 1$ to the previous result. Since 1 is also $2 \times 0 + 1$, also $\mathsf{myFun}(0)$ has the same shape. In conclusion, we can write:

$$\mathsf{myFun}(n) = \sum_{i=0}^{n} (2i + 1).$$

In general, if we already have a sequence of numbers a_0, a_1, a_2, and so on, we can add them up and the result of the addition is denoted by $\sum_{i=0}^{n} a_i$. As a recursive function, this summation can be defined as follows:

$$\begin{aligned}
&\sum_{i=0}^{0} a_i = a_0 \\
&\sum_{i=0}^{n} a_i = \left(\sum_{i=0}^{n-1} a_i \right) + a_n \qquad \text{if } n > 0.
\end{aligned}$$

The starting point of a summation doesn't have to be zero. We can choose to start adding at any natural value. So, for example, we can compute

$$\begin{aligned}
\sum_{i=3}^{6} (3 \times i - 5) &= (3 \times 3 - 5) + (3 \times 4 - 5) + (3 \times 5 - 5) + (3 \times 6 - 5) \\
&= 4 + 7 + 10 + 13 = 34.
\end{aligned}$$

Nor need it be a natural number: it is allowed to compute summations with bounds that are integers. So it is possible to start summing with a negative index, as in this example:

$$\begin{aligned}\sum_{i=-2}^{2} i^2 &= (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 \\ &= 4 + 1 + 0 + 1 + 4 = 10.\end{aligned}$$

Convention. Obviously, it doesn't make any sense to compute a summation in which the starting index is larger than the final one. What would it mean to compute

$$\sum_{i=10}^{5} i \quad ?$$

This kind of summations, however, sometimes pop up when proving mathematical properties of programs. So it is a good idea to give them a conventional value. Since we are supposed to stop summing before we even begin, it makes sense to set the value of such a summation to zero. So, in general:

$$\sum_{i=n}^{m} a_i = 0 \qquad \text{if } m < n.$$

Splitting Summations. A summation can be split in two parts by separating the range over which we do the addition in two sections. For example:

$$\sum_{i=0}^{10} a_i = \sum_{i=0}^{5} a_i + \sum_{i=6}^{10} a_i$$

In general, we can split a summation at any point k inside its range:

$$\sum_{i=n}^{m} a_i = \sum_{i=n}^{k} a_i + \sum_{i=k+1}^{m} a_i \qquad \text{if } n \leq k < m.$$

The simplest form of splitting happens when we just extract the last term from the summation:

$$\sum_{i=n}^{m} a_i = \sum_{i=n}^{m-1} a_i + a_m \qquad \text{if } n \leq m.$$

3.8 Induction

The way the natural numbers are generated helps us not only in defining functions, but also in proving properties. If we want to prove that a certain statement about numbers is true for every element of $\mathbb{N}$, we can do it by showing that it is true for 0 and that if it is true for a number k, then it must be true for the next number $k+1$.

This method to prove propositions about natural numbers is called *induction*. Suppose that the property you're interested in is called P. To say that a number n satisfies this property, we write $P(n)$.

$$\textbf{Induction Principle:}\quad \begin{array}{rl} \text{if} & P(0) \\ \text{and} & P(k) \Rightarrow P(k+1) \text{ for all numbers } k \\ \hline \text{then} & P(n) \textbf{ for all natural numbers } n. \end{array}$$

The idea is that we can prove that P holds for all natural numbers one step at a time:

- $P(0)$ holds by hypothesis;
- Instantiating the step case with $k = 0$ we get that $P(0) \Rightarrow P(1)$ and so, since we already know that $P(0)$ is true from the previous step, also $P(1)$ is true. (This is the rule of implication elimination, if you remember well.)
- Instantiating the step case with $k = 1$ we get that $P(1) \Rightarrow P(2)$ and so, since we already know that $P(1)$ is true from the previous step, also $P(2)$ is true.
- With $k = 2$ we get that $P(2) \Rightarrow P(3)$, and $P(2)$ is true by the previous step, so also $P(3)$ is true.
- With $k = 3$ we get that $P(3) \Rightarrow P(4)$, and $P(3)$ is true by the previous step, so also $P(4)$ is true.
- And so on for all the natural numbers.

The two cases of the proof by induction are called the **base case** ($P(0)$) and the **induction case** ($P(k) \Rightarrow P(k+1)$). When proving the induction case, we make the assumption $P(k)$, which is called the **induction hypothesis**, and from it we must prove the conclusion $P(k+1)$.

Example 1. Let's choose a simple property to show how induction works. Set $P(n)$ to be the following proposition:

$$P(n): \quad n^2 + 1 > n.$$

To prove that $P(n)$ is true for all natural numbers n, we need to prove the two hypotheses of the induction principle:

$P(0)$**:** When we replace n with 0 in the formula for $P(n)$ we get: $0^2 + 1 > 0$, that is, $1 > 0$, which is trivially true.

$P(k) \Rightarrow P(k+1)$ **for every** k**:** To show this, we can assume that $P(k)$ is true and, using this assumption, we must be able to prove that also $P(k+1)$ is true.

$$\textbf{Induction Hypothesis } P(k): k^2 + 1 > k.$$

Using the assumption above we must prove:

$$\textbf{Conclusion to prove } P(k+1) : (k+1)^2 + 1 > k+1.$$

In fact this can be derived:

$$\begin{aligned}(k+1)^2+1 &= k^2+2k+1+1 = k^2+1+2k+1 \\ &> k+2k+1 \qquad \text{by the induction hypothesis } (k^2+1>k) \\ &\geq k+1.\end{aligned}$$

Since we managed to prove both $P(0)$ and $P(k) \Rightarrow P(k+1)$, we can conclude that $P(n)$ must be true for every natural number n. That is, for every number n, we know that $n^2+1>n$.

Example 2. This time let's choose the following property of natural numbers:

$$P(n): \quad 7^n - 1 \text{ is divisible by } 6.$$

We can check that this is true by the first few natural numbers, $0, 1, 2, 3$, by computing the formula for each of them: $7^0-1=0$, $7^1-1=6$, $7^2-1=48$, $7^3-1=342$; each of these numbers is a multiple of 6, so the property is satisfied up to $n=3$. But if we want to prove that it is always true for all natural numbers, we must use induction. So we need to prove the base case and the inductive step:

$P(0)$**:** We must prove that 7^0-1 is divisible by 6. But $7^0-1=0$ and 0 is a multiple of every number. So $P(0)$ is true.

$P(k) \Rightarrow P(k+1)$ **for every** k**:** To show this, we can assume that $P(k)$ is true and, using this assumption, we must be able to prove that also $P(k+1)$ is true.

$$\textbf{Induction Hypothesis } P(k) : 7^k - 1 \text{ is divisible by } 6.$$

That is, there is some number h such that $7^k - 1 = h \times 6$. Using the assumption above we must prove:

$$\textbf{Conclusion to prove } P(k+1) : 7^{k+1} - 1 \text{ is divisible by } 6.$$

In fact we can make the following derivation:

$$\begin{aligned}7^{k+1}-1 &= 7 \times 7^k - 1 = 7 \times (7^k - 1 + 1) - 1 \\ &= 7 \times (h \times 6 + 1) - 1 \qquad \text{by the induction hypothesis } 7^k - 1 = h \times 6 \\ &= 7 \times h \times 6 + 7 - 1 = (7 \times h + 1) \times 6\end{aligned}$$

which shows that also $7^{k+1}-1$ is a multiple of 6, that is, that $P(k+1)$ is true.

Since we proved both $P(0)$ and $P(k) \Rightarrow P(k+1)$, we can conclude that $P(n)$ must be true for every natural number n. That is, for every number n, we know that $7^n - 1$ is divisible by 6.

Example 3. As our last example, we prove a *closed expression* for the addition of the first n numbers:

$$P(n): \quad \sum_{i=0}^{n} i = \frac{n(n+1)}{2}.$$

Let's check that this is indeed the case for the first few values of n:

$$\begin{aligned}
\sum_{i=0}^{0} i &= & 0 &= \frac{0(0+1)}{2} \\
\sum_{i=0}^{1} i &= & 0+1=1 &= \frac{1(1+1)}{2} \\
\sum_{i=0}^{2} i &= & 0+1+2=3 &= \frac{2(2+1)}{2} \\
\sum_{i=0}^{3} i &= & 0+1+2+3=6 &= \frac{3(3+1)}{2} \\
\sum_{i=0}^{4} i &= & 0+1+2+3+4=10 &= \frac{4(4+1)}{2}
\end{aligned}$$

So the equality seems to hold. To be absolutely sure that it is always true, though, we need to give a proof by induction.

$P(0)$**:** We must prove that $\sum_{i=0}^{0} i = 0(0+1)/2$. We just verified this equality in the first line above. So $P(0)$ is true.

$P(k) \Rightarrow P(k+1)$ **for every** k**:** To show this, we can assume that $P(k)$ is true and, using this assumption, we must be able to prove that also $P(k+1)$ is true.

$$\textbf{Induction Hypothesis } P(k): \sum_{i=0}^{k} i = \frac{k(k+1)}{2}.$$

Using the assumption above we must prove:

$$\textbf{Conclusion to prove } P(k+1): \sum_{i=0}^{k+1} i = \frac{(k+1)(k+2)}{2}.$$

In fact we can extract the last term of the summation, then apply the assumption and after some simple arithmetical manipulation we get the desired result:

$$\begin{aligned}\sum_{i=0}^{k+1} i = \sum_{i=0}^{k} i + (k+1) & \\ = \frac{k(k+1)}{2} + (k+1) & \qquad \text{by the induction hypothesis} \\ = \frac{k(k+1)+2(k+1)}{2} = \frac{(k+1)(k+2)}{2}. &\end{aligned}$$

which shows that also $P(k+1)$ is true.

Since we proved both $P(0)$ and $P(k) \Rightarrow P(k+1)$, we can conclude that $P(n)$ must be true for every natural number n. That is, for every number n, we know that $\sum_{i=0}^{n} i = n(n+1)/2$.

3.9 The Axioms of Natural Numbers

The fundamental properties of the set $\mathbb{N}$ of natural numbers can be summarized by a sequence of axioms and definitions. An *axiom* is a proposition that is simple enough so that we may consider it obviously true. We assume the axioms to be true without the need of any proof. Every other assertion about the natural numbers must be proved. A proposition that can be derived from the axioms and definitions using logical rules is called a *theorem.*

Addition:

$a + 0 = a$	zero is the identity of addition
$a + b = b + a$	commutativity of addition
$(a + b) + c = a + (b + c)$	associativity of addition
$a + c = b + c \Rightarrow a = b$	cancellation for addition
$a + 1 \neq 0$	injectivity of successor

Multiplication:

$a \cdot 0 = 0$	zero absorbs multiplication
$a \cdot 1 = a$	one is the identity of multiplication
$a \cdot b = b \cdot a$	commutativity of multiplication
$(a \cdot b) \cdot c = a \cdot (b \cdot c)$	associativity of multiplication
$a \cdot c = b \cdot c \wedge c \neq 0 \Rightarrow a = b$	cancellation for multiplication
$a \cdot (b + c) = a \cdot b + a \cdot c$	distributivity of mult. over add.

Order. Definitions of the four order relations:

$$\begin{array}{lcl} a \le b & \Leftrightarrow & b = a + h \quad \text{for some natural number } h \\ a < b & \Leftrightarrow & a + 1 \le b \\ a \ge b & \Leftrightarrow & b \le a \\ a > b & \Leftrightarrow & b < a \end{array}$$

Principle of Induction. For every property P of natural numbers:

$$\begin{array}{rl} \text{if} & P(0) \\ \text{and} & P(k) \Rightarrow P(k+1) \text{ for all numbers } k \\ \hline \text{then} & P(n) \textbf{ for all natural numbers } n. \end{array}$$

3.10 Derived Theorems

All other properties of natural numbers should be derived from the definitions and axioms using the induction principle and the rules of logic. For example, earlier we showed these three properties of order:

$$\begin{array}{ll} a \le a & \text{Reflexivity} \\ a \le b \wedge b \le a \Rightarrow a = b & \text{Antisymmetry} \\ a \le b \wedge b \le c \Rightarrow a \le c & \text{Transitivity} \end{array}$$

Before we go on and prove more theorems, let me explain how the axioms and definitions are used.

Instantiation of Axioms and Definition. Each of the axioms and definitions uses variable names, $a, b, c, \ldots, x, y, z$. The meaning is that the axiom is true for all possible values of those variables. Therefore, whenever we use an axiom, we are allowed to substitute all the variables with arbitrary expressions. Here is an example, instantiating the distributivity axiom.

$$\begin{array}{lll} a \cdot (b + c) = a \cdot b + a \cdot c & \text{distributivity of mult. over add.} & \\ (x+1) \cdot (y+2) = (x+1) \cdot y + (x+1) \cdot 2 & \text{substitute:} & a \mapsto x + 1 \\ & & b \mapsto y \\ & & c \mapsto 2 \end{array}$$

Be careful: if you substitute not just a value or a variable, but a more complex expression, like $x+1$ in the example, you should put parentheses around it when you write it in the formula.

Substitution of Equality. Another important principle is that of *substitution of equal terms*. This means that, if you know that two expressions are equal, then you can replace one for the other in any formula. Suppose $\mathcal{C}[\cdots]$ is a formula or expression containing a *hole*, here indicated by $\cdots$, that can be filled by a term. Then we have the principle:

$$a = b \Rightarrow \mathcal{C}[a] = \mathcal{C}[b]$$

That is: if a and b are equal, then the formula obtained by placing a in the hole is equal to the formula obtained by placing b in the hole. Another way of reading the conclusion is: an expression containing a is equal to the expression obtained by replacing some of the occurrences of a by b.

Here is an example of the use of this principle:

$$(x+1)^2 = x^2 + 2 \cdot x + 1 \qquad \textit{can you prove this from the axioms?}$$

$$\frac{(x+1)^2}{7+(x+1)^2} = \frac{x^2+2\cdot x+1}{7+(x^2+2\cdot x+1)} \qquad \text{substitution of equality}$$

In this case, the role of a is played by $(x+1)^2$, the role of b is played by $x^2 + 2 \cdot x + 1$, and the formula with holes $\mathcal{C}[\cdots]$ is $\frac{\cdots}{7+\cdots}$.

Let's look at a couple of simple theorem to get a feeling of how propositions are derived from the axioms. We are going to show that our definition of the strict order relation $a < b$ is equivalent to saying that $a \leq b$ and $a \neq b$. Let's prove each part separately.

Theorem 6

$$a < b \Rightarrow a \leq b.$$

Proof. Assume that $a < b$ is true.

By definition of the strict order relation, this means that $a+1 \leq b$ and by definition of the non-strict order relation, this means that there exists a number h such that $(a+1)+h = b$. But then we have that $a+(h+1) = b$, which, again by definition of the non-strict relation, implies that $a \leq b$.

So, from the assumption $a < b$ we derived the conclusion $a \leq b$. This means that we proved the statement of the theorem: $a < b \Rightarrow a \leq b$. □

Theorem 7

$$a < b \Rightarrow a \neq b.$$

Proof. Assume that $a < b$.

As before, this means that $a+1 \leq b$. We need now to derive $a \neq b$: we must show that $a = b$ cannot be true. We assume that $a = b$ is true and show that this leads to a contradiction.

$a+1 = b+1$	substitution of equality in $\cdots + 1$
$b \leq a+1$	definition of $\leq$
$b = a+1$	antisymmetry from $a+1 \leq b$ and $b \leq a+1$
$a = a+1$	substitute $a = b$ in $b = a+1$

This last statement is clearly impossible *(exercise: prove rigorously that it is false)*. Since our assumption $a = b$ leads to a contradiction, we can conclude that $a \neq b$. □

Notice that we could turn this proof, written in English, into a derivation in formal logic:

$$\begin{array}{lll}
1 & \mid \; \mid \; \underline{a < b} & \\
2 & \mid \; \mid \; \mid \; \underline{a = b} & \\
\vdots & \mid \; \mid \; \mid \; \vdots & \text{arithmetic steps} \\
k & \mid \; \mid \; \mid \; \bot & \\
(k+1) & \mid \; \mid \; a \neq b & \neg\text{I, } 2\text{–}k \\
(k+2) & \mid \; a < b \Rightarrow a \neq b & \Rightarrow\text{I, } 1\text{–}(k+1)
\end{array}$$

In principle, every rigorous proof can be turned into such a derivation. However, for the time being we will be satisfied with more informal proofs written in English.

As an exercise, try to prove the inverse implication by yourself:

$$a \leq b \wedge a \neq b \Rightarrow a < b.$$

The rules of algebra are supposed to be familiar to everybody from high-school. Therefore, we allow ourselves to manipulate expressions and prove equalities of expressions without justifying every single step by the axioms. We will just write *this is true by simple arithmetic.*

Consider this strengthened version of the transitivity property.

Theorem 8

$$a \leq b \wedge b < c \Rightarrow a < c.$$

Proof. Assume that $a \leq b$ and $b < c$. The first assumption is equivalent to $a + h = b$ for some number h, by definition of $\leq$. The second assumption is equivalent to $b + 1 \leq c$ by definition of $<$ and therefore it is equivalent to $(b + 1) + k = c$ for some number k, by definition of $\leq$.

Then we can prove the following equality:

$$\begin{array}{rll}
c & = b + 1 + k & \text{by the second assumption} \\
 & = (a + h) + 1 + k & \text{by the first assumption (substitution)} \\
 & = (a + 1) + (h + k) & \text{simple arithmetic}
\end{array}$$

We proved that $(a + 1) + (h + k) = c$, which tells us, by definition of $\leq$, that $a + 1 \leq c$, that is, by definition of $<$, that $a < c$. □

This theorem, together with transitivity of $\leq$ and substitution of equality, justifies using a chain of equalities and inequalities to prove a certain goal. From a chain like this:

$$\begin{array}{ll}
\textsf{initial expression} & \leq \textsf{expression}_1 \\
 & = \textsf{expression}_2 \\
 & < \textsf{expression}_3 \\
 & \leq \textsf{final expression}
\end{array}$$

we can conclude that: $\mathsf{initial\ expression} < \mathsf{final\ expression}$.

If no strict order was part of the chain, also the conclusion will be non-strict. From:

$$\begin{aligned}\mathsf{initial\ expression} \quad &\leq \mathsf{expression}_1\\ &= \mathsf{expression}_2\\ &\leq \mathsf{expression}_3\\ &= \mathsf{final\ expression}\end{aligned}$$

we can conclude that: $\mathsf{initial\ expression} \leq \mathsf{final\ expression}$.