Math 220 Fall 2021, Lecture 14: Mathematical induction

1. Mathematical induction (cont'd)

1.1. What is induction? (cont'd)

Last time, we stated (but didn't get around to prove) the following theorem:

Theorem 1.1.1. Let x and y be any two numbers. Then, for any $n \in \mathbb{N}$, we have

$$(x-y)\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}\right)=x^n-y^n.$$

The big sum in the parentheses is the sum of all products of the form x^iy^j where i and j are nonnegative integers with $i+j=n-1$.

Proof. We induct on n.

Base case: For $n=0$, the claim of the theorem is

$$(x-y)\underbrace{\left(x^{0-1}+x^{0-2}y+x^{0-3}y^2+\cdots+x^2y^{0-3}+xy^{0-2}+y^{0-1}\right)}_{\substack{\text{This should be understood as the sum of all products of the form } x^iy^j\\ \text{where } i \text{ and } j \text{ are nonnegative integers with } i+j=0-1.\\ \text{Since there are no such products, this sum is empty, and thus equals } 0.}}=x^0-y^0,$$

which is easily checked (keep in mind that $x^0=1$ and $y^0=1$, so the right hand side is $1-1=0$).

Induction step: Let $n \in \mathbb{N}$. Assume that the theorem holds for n. We must prove that it holds for $n+1$.

We have assumed that

$$(x-y)\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}\right)=x^n-y^n. \tag{1}$$

We must prove that

$$(x-y)\left(x^{n}+x^{n-1}y+x^{n-2}y^2+\cdots+x^2y^{n-2}+xy^{n-1}+y^{n}\right)=x^{n+1}-y^{n+1}.$$

We have

$$\begin{aligned}
&(x-y)\left(\underbrace{x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + x^2y^{n-2} + xy^{n-1}}_{=x\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+xy^{n-2}+y^{n-1}\right)} + y^n\right) \\
&= (x-y)\left(x\left(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1}\right) + y^n\right) \\
&= (x-y)\,x\left(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1}\right) + (x-y)\,y^n \\
&= x\underbrace{(x-y)\left(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1}\right)}_{\substack{=x^n-y^n \\ \text{(by the induction hypothesis (1))}}} + (x-y)\,y^n \\
&= x\left(x^n - y^n\right) + (x-y)\,y^n = x^{n+1} - xy^n + xy^n - y^{n+1} = x^{n+1} - y^{n+1}.
\end{aligned}$$

This is precisely what we needed to prove. So the induction step is complete, and the theorem is proved. □

In the future, we will often encounter sums like

$$x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + x^2y^{n-3} + xy^{n-2} + y^{n-1},$$

which take a while to decipher. There is a notation that makes such sums shorter and easier to understand. This is the **sigma notation** (aka the **summation sign**). In its simplest form, it is defined as follows:

Definition 1.1.2. Let u and v be two integers. Let $a_u, a_{u+1}, \ldots, a_v$ be any numbers. Then,

$$\sum_{k=u}^{v} a_k$$

is defined to be the sum

$$a_u + a_{u+1} + \cdots + a_v.$$

(If $v < u$, then this sum is understood to be 0.)

For example:

$$\begin{aligned}
&\sum_{k=5}^{10} k = 5 + 6 + 7 + 8 + 9 + 10 = 45; \\
&\sum_{k=5}^{10} \frac{1}{k} = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} = \frac{2131}{2520}; \\
&\sum_{k=5}^{10} k^k = 5^5 + 6^6 + 7^7 + 8^8 + 9^9 + 10^{10}.
\end{aligned}$$

The variable k is not set in stone; it is just supposed to be a variable that is not used for anything else. So, for example, you can rewrite $\sum\limits_{k=u}^{v} a_k$ as $\sum\limits_{\ell=u}^{v} a_\ell$ or $\sum\limits_{x=u}^{v} a_x$ or $\sum\limits_{\spadesuit=u}^{v} a_\spadesuit$. Just don't use a variable that already has a meaning. For example, you cannot write $\sum\limits_{u=u}^{v} a_u$.

A couple more examples, this time with an "outside variable". So let n be an arbitrary nonnegative integer. Then,

$$\sum_{k=1}^{n} k = 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \qquad \text{(by a theorem from Lecture 12)};$$

$$\begin{aligned}\sum_{k=1}^{n} (2k-1) &= (2 \cdot 1 - 1) + (2 \cdot 2 - 1) + \cdots + (2 \cdot n - 1) \\ &= 1 + 3 + \cdots + (2n-1) = n^2 \qquad \text{(by an exercise on homework set \#2)};\end{aligned}$$

$$\sum_{k=0}^{n-1} 2^k = 2^0 + 2^1 + \cdots + 2^{n-1} = 2^n - 1 \qquad \text{(by a theorem from Lecture 13)};$$

$$\begin{aligned}\sum_{k=1}^{n} k^2 &= 1^2 + 2^2 + \cdots + n^2 \\ &= \frac{n(n+1)(2n+1)}{6} \qquad \text{(this can, too, be proved by induction)}.\end{aligned}$$

Not every sum like this has a closed-form expression. For instance, there is no closed form for

$$\sum_{k=1}^{n} \frac{1}{k} \qquad \text{or for} \qquad \sum_{k=1}^{n} k^k.$$

(By "closed form", I mean an expression that involves no $\sum$ signs and no "$\cdots$"s.)

The sum

$$x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + x^2y^{n-3} + xy^{n-2} + y^{n-1}$$

in our first theorem today can now be rewritten as

$$\sum_{k=0}^{n-1} x^{n-1-k}y^k.$$

So the claim of the theorem rewrites as

$$(x-y)\left(\sum_{k=0}^{n-1} x^{n-1-k}y^k\right) = x^n - y^n.$$

The notation $\sum\limits_{k=u}^{v} a_k$ is called **sigma notation** or **finite sum notation**. The symbol $\sum$ is called the **summation sign**. The numbers u and v are called the **lower limit** and the **upper limit** of the summation; the variable k is called the **summation index** or the **running index**. The sum $\sum_{k=u}^{v} a_k$ is called the **sum of the a_k for k running** (or **ranging**) **from u to v**.

Similarly to the finite sum notation, there is a **finite product notation**:

$$\prod_{k=u}^{v} a_k := a_u a_{u+1} \cdots a_v.$$

The symbol $\prod$ is called the **product sign**. The k is called the **product index**. Here are some examples:

$$\prod_{k=5}^{10} k = 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 = 151\,200;$$

$$\prod_{k=1}^{n} k = 1 \cdot 2 \cdot \cdots \cdot n = n! \qquad \text{(as defined on homework set \#2)};$$

$$\prod_{k=1}^{n} a = \underbrace{aa \cdots a}_{n \text{ times}} = a^n \qquad \text{for any fixed number } a.$$

Just as an empty sum is defined to be 0, an empty product (i.e., a product with no factors) is defined to be 1. Therefore, $a^0 = 1$ for any number a. (In particular, $0^0 = 1$.) For the same reason, $0! = 1$.

Here is another example:

$$\begin{aligned}
\prod_{k=1}^{n} 2^k &= 2^1 \cdot 2^2 \cdot 2^3 \cdot \cdots \cdot 2^n \\
&= 2^{1+2+3+\cdots+n} && \left(\begin{array}{c}\text{by one of the laws of exponents: namely,} \\ \text{the law saying } a^{i_1} a^{i_2} \cdots a^{i_k} = a^{i_1+i_2+\cdots+i_k}\end{array}\right) \\
&= 2^{n(n+1)/2} && \text{(since we know that } 1+2+3+\cdots+n = n\,(n+1)\,/2\text{)}.
\end{aligned}$$

Let us now see a more combinatorial way of using induction.

Let us analyze the **Tower of Hanoi puzzle**. This is a puzzle that involves three "towers" (pegs). Two of the towers are empty; the third one has n disks stacked one upon the other. The disks have sizes $1, 2, \ldots, n$, and are stacked in the order of decreasing size (the largest disk being at the very bottom). You are required to move all disks to another tower. You are allowed to make moves of the form "move the topmost disk from one tower to another", but you cannot do such a move if this would lead to a larger disk being placed on top of a smaller one. (You can play this game at `https://www.mathsisfun.com/games/towerofhanoi.html` .)

Theorem 1.1.3. The puzzle can always be solved in $2^n - 1$ moves.

Proof. We induct on n:

Base case: For $n = 0$, the puzzle can be solved in $2^0 - 1$ moves, because there is nothing to move (and $2^0 - 1 = 0$).

Induction step: Let $n \in \mathbb{N}$. Assume (as the induction hypothesis) that the puzzle with n disks can be solved in $2^n - 1$ moves. We must prove that the puzzle with $n + 1$ disks can be solved in $2^{n+1} - 1$ moves.

We prove this as follows: Consider the puzzle with $n + 1$ disks. The biggest of the disks lies at the very bottom of its tower; we will call it the **big disk**. Now, we apply our induction hypothesis to move the n other disks to another tower; this requires $2^n - 1$ moves. (The big disk does not get in our way, because any disk can be stacked on top of it.) Next, we move the big disk to the free tower. Finally, we again apply the induction hypothesis to move the n other disks to the tower that has the big disk. (Again, the big disk does not give us any trouble.) Now, all the disks are on that tower. In total, this playthrough has required

$$(2^n - 1) + 1 + (2^n - 1) = \underbrace{2 \cdot 2^n}_{=2^{n+1}} - 1 = 2^{n+1} - 1$$

many moves. Thus, we have moved all the disks to a different tower in $2^{n+1} - 1$ many moves. This completes the induction step, and thus the proof of the theorem. $\square$

Question 1.1.4. Can the puzzle be solved in fewer than $2^n - 1$ moves?