Induction

David Arthur
darthur@gmail.com

1 Induction

I think you have probably all seen induction. It's pretty great, and we're going to be practicing it today. Mostly I'll be talking through a few examples and giving problems, but let's start with a few tips up-front.

> **Remember to try it:**
>
> The hardest part of an induction problem is often remembering to use induction! When solving problems, *always* consider induction!

Here are some situations where induction is especially helpful:

- Any problem involving a recurrence relationship or a sequence.
- Any problem involving counting, because you can often reduce a counting problem to a recurrence relationship.
- Any problem involving games. If you can guess exactly what the winning positions are, it's probably easy to prove your guess with induction.
- Functional equations, especially over the integers or rational numbers. Can you calculate $f(nx)$ in terms of $f(x)$ for example? That's always useful!
- Proving something exists. You can do this by inductively assuming something smaller exists, and then extending it to what you want.

> **Tweak the inductive hypothesis:**
>
> When you try to prove an induction statement, you may find that your inductive hypothesis is not giving you all the information you want. If that happens, remember that you can change the inductive hypothesis! The inductive hypothesis doesn't have to be exactly what the problem gives you. Just ask yourself: what do you really wish you could assume about the smaller case?

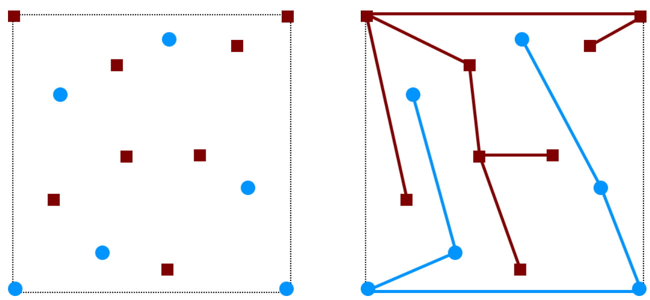
A classic example of this is proving $\sum_{n=1}^{\infty} \frac{1}{(n+1)\sqrt{n}} < 2$. Letting $S_N = \sum_{n=1}^{N} \frac{1}{(n+1)\sqrt{n}}$, we would want to prove an upper bound for S_N by induction. If all you know is $S_{N-1} < 2$, then all you can get is $S_N < 2 + \frac{1}{(n+1)\sqrt{n}}$, which isn't strong enough. On the other hand, you could instead prove by induction that $S_N < 2 - \frac{2}{\sqrt{N+1}}$, and then everything will just work! Just don't forget to check the base case too.

> **Be smart in how you reduce to smaller cases:**
>
> Induction lets you understand *all* smaller cases. In school, you often get taught to just look at $n-1$ in the most obvious way. But you should be creative – think about different ways of reducing to smaller cases and see what is most helpful.

This one is a bit tougher, but you'll get if you practice.

2 Examples

Let's start with a cute problem from the 2005 International Olympiad in Informatics that uses the principle of "tweaking the inductive hypothesis".



Example 1. *A number of red points and blue points are drawn in a unit square with the following properties:*

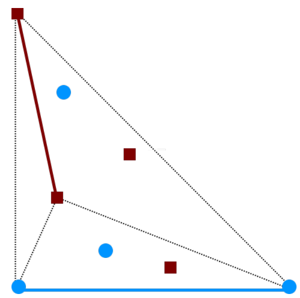
- *The top-left and top-right corners are red points.*
- *The bottom-left and bottom-right corners are blue points.*
- *No three points are collinear.*

Prove it is possible to draw red segments between red points and blue segments between blue points in such a way that: all the red points are connected to each other, all the blue points are connected to each other, and no two segments cross.[1]

Solution. This question is just being mean. It's really tricky to see how to make progress if the points are in a square, but let's instead focus on the bottom left *triangle*. We will prove the following:

> *Consider a triangle where two vertices are one colour and already connected by a segment, and the third vertex is the other colour. Then it is possible to draw segments inside the triangle connecting all points of the same colour.*

[1] International Olympiad in Informatics 2006



We will prove this by induction on the number of points inside the triangle. If the triangle has no points inside, it's obvious.

Otherwise consider a triangle ABC, and assume without loss of generality that AB is blue and C is red. If ABC has no internal red points, we can just connect everything with blue segments and we're done. Otherwise, ABC has some internal blue point P. We will colour CP red. Each triangle ABP, BCP, CAP has fewer internal points than ABC (since P is not inside any of them), so by the inductive hypothesis, it is possible to draw segments inside each triangle ABP, BCP, CAP to connect all points of the same colour. But then every red point in ABC is connected to P, and every blue point is connected to *either* A or B and those two points are directly connected to each other. Thus, we can also connect ABC and the inductive result follows.

Therefore, we can connect all the red and blue points in both the bottom-left triangle and the upper-right triangle of our original picture, and we're done! $\square$

You may have already seen the next problem if you read notes online. It appeared on the 1999 APMO, and it is a nice example of the "be smart in how you reduce to smaller cases" principle.

Example 2. *Let $a_1, a_2, \ldots, a_n$ be a sequence of real numbers satisfying $a_{i+j} \le a_i + a_j$ for all positive integers i, j. Prove that*

$$a_1 + \frac{a_2}{2} + \ldots + \frac{a_n}{n} \ge a_n$$

for each positive integer n.

Solution. Certainly the result holds for $n = 1$. Now suppose the result holds for values less than n, and we will try to show it for n. We know $a_1 + \frac{a_2}{2} + \frac{a_3}{3} + \ldots + \frac{a_{n-1}}{n-1} \ge a_{n-1}$, so the obvious thing to try is to show $a_{n-1} + \frac{a_n}{n} \ge a_n$.

Unfortunately this turns out not to be true. For example, if $(a_1, a_2, a_3) = (1, 1, 2)$, you can check the conditions of the problem are satisfied but $a_2 + \frac{a_3}{3} = \frac{5}{3} < 2 = a_3$. Oops! It is tempting to give up on induction at this point, but all is not lost!

Using the result for each value less than n, we have:

$$\begin{aligned} a_1 &\ge a_1 \\ a_1 + \frac{a_2}{2} &\ge a_2 \\ &\vdots \\ a_1 + \frac{a_2}{2} + \ldots + \frac{a_{n-1}}{n-1} &\ge a_{n-1} \end{aligned}$$

If we add these all up, we get

$$\begin{aligned}(n-1)a_1 + \frac{(n-2)a_2}{2} + \ldots + \frac{a_{n-1}}{n-1} &\geq a_1 + a_2 + \ldots + a_{n-1} \\ \implies na_1 + \frac{na_2}{2} + \ldots + \frac{na_{n-1}}{n-1} &\geq 2(a_1 + a_2 + \ldots + a_{n_1})\end{aligned}$$

However, the problem condition directly gives us $2(a_1 + a_2 + \ldots + a_{n-1}) \geq (n-1)a_n$, so we are now done. $\square$

Finally, I wanted to also give a functional equation example. These are pretty different from other problems and yet very important on Olympiads!

Example 3. *Let $\mathbb{Q}$ be the set of rational numbers. Find all functions f, defined on $\mathbb{Q}$ and taking values in $\mathbb{Q}$, such that*

$$f(x+y) + f(y+z) + f(z+x) + f(0) = f(x+y+z) + f(x) + f(y) + f(z)$$

for all rational numbers x, y, z.

Solution. First suppose $f(x) = Ax^2 + Bx + C$ for some rational numbers A, B, C. Then

$$\begin{aligned}& f(x+y) + f(y+z) + f(z+x) + f(0) \\ = \;& A \cdot \Big((x+y)^2 + (y+z)^2 + (z+x)^2\Big) + B \cdot \Big((x+y) + (y+z) + (z+x)\Big) + 4C \\ = \;& A \cdot \Big((x+y+z)^2 + x^2 + y^2 + z^2\Big) + B \cdot \Big((x+y+z) + x + y + z\Big) + 4C \\ = \;& f(x+y+z) + f(x) + f(y) + f(z)\end{aligned}$$

Thus, any quadratic polynomial satisfies the given equation.

Conversely, suppose f satisfies the given equation and fix a rational number U. We claim there exists a quadratic polynomial P such that $f(x) = P(x)$ for $x \in \{-U, 0, U\}$. Indeed, the following polynomial works:

$$P(x) = \frac{(x)(x-u)}{2u^2} \cdot f(-u) + \frac{(x-u)(x+u)}{-u^2} \cdot f(0) + \frac{(x)(x+u)}{2u^2} \cdot f(u).$$

Now suppose $f(x) = P(x)$ for $x \in \{-U, 0, U, \ldots, (n-1)U\}$ for some $n \geq 2$ (*). We can then let $x = (n-1)U, y = U, z = -U$ in the given functional equation to find:

$$\begin{aligned}f(nU) + f(0) + f\big((n-2)U\big) + f(0) &= f\big((n-1)U\big) + f\big((n-1)U\big) + f(U) + f(-U) \\ \implies f(nU) &= 2f\big((n-1)U\big) - f\big((n-2)U\big) + f(U) + f(-U) - 2f(0) \\ &= 2P\big((n-1)U\big) - P\big((n-2)U\big) + P(U) + P(-U) - 2P(0).\end{aligned}$$

The last step here came from applying (*). However, we already showed above that $P(x)$ also satisfies the functional equation, and therefore $P(nU)$ equals the same thing. In other words, we have shown that if $f(x) = P(x)$ for $x \in \{-U, 0, U, \ldots, (n-1)U\}$, then $f(nU) = P(nU)$ as well. It follows from induction that $f(nU) = P(nU)$ for *all* positive integers n. And if we replace U with $-U$, we also obtain $f(nU) = P(nU)$ for all negative integers n.

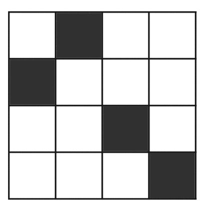
If we take $U = 1$, this proves there is a quadratic polynomial P such that $f(x) = P(x)$ for all integers. But what about rational numbers? Consider an arbitrary rational number $\frac{p}{q}$ and let $U = \frac{1}{q}$. As argued above, there is a quadratic polynomial P' such that $f\left(\frac{n}{q}\right) = P'\left(\frac{n}{q}\right)$ for all integers n. Taking $n = 0, q, 2q$, we have $P'(x) = f(x) = P(x)$ for $x = 0, 1, 2$. However, two distinct quadratic polynomials can share a value at only two points. Therefore, $P'(x)$ must be identical to $P(x)$, and hence $f\left(\frac{p}{q}\right) = P'\left(\frac{p}{q}\right) = P\left(\frac{p}{q}\right)$ for all rational numbers $\frac{p}{q}$.

□

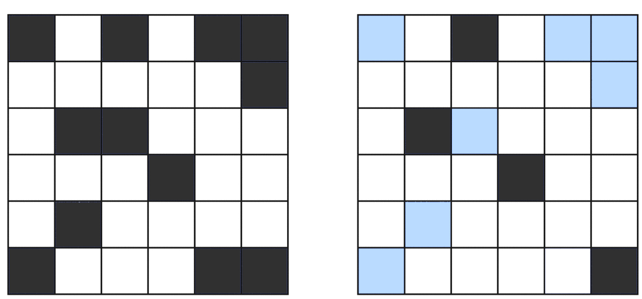
3 A Real Research Example

One of the great things about Olympiads is the techniques stay with you even as you go on in mathematics. In this section, I'm going to talk about an actual math research paper that uses induction. It is a short proof to a conjecture that had remained unsolved for 15+ years, and one of the two authors was a student only just beginning graduate school! It is called the Stanley-Wilf conjecture, and was proven in 2004 by Adam Marcus and Gábor Tardos. This is just for fun so it's okay if you don't follow it all, but I hope you find it interesting.

Let P be a fixed "pattern": a $k \times k$ grid consisting of white and black squares with at most one black square in each row and column. See for example here:



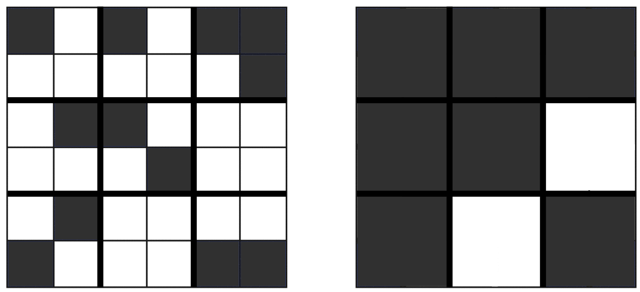
Consider a larger grid G, also consisting of white and black squares. We are interested in whether it has k squares that form the same shape as P. For example, consider the grid G on the left here. If you just look at just the four squares highlighted to the right, you can see they are in the same shape as P:



More precisely, we will say that G *contains the pattern* P if it is possible to delete rows and columns from G until it is the same size as P, and then have black squares remaining in every position that G has a black square. Otherwise we say G *avoids the pattern* P.

Theorem 1. *Fix a pattern P, and let X_n denote the maximum number of black squares in an $n \times n$ grid that* avoids *the pattern P. Then there exists a number C such that $X_n \leq C \cdot n$ for all n.*

Consider some $n \times n$ grid G that avoids P. To apply induction, we need to reduce the problem to a smaller size, which we'll do by dividing G into "blocks" of size $m \times m$ for some m. We will say that a block is "white" if every square in it is white, and we'll say the block is black otherwise. The blocks then describe a grid G' of size $\frac{n}{m}$. For example, if we divide our example grid from above into 2×2 blocks, it would look like the following:



The first key observation is that G' avoids P. If it did not, there is a set of k blocks in the same shape as P. If we choose a square from each block, then those squares will also be in the same shape as P, contradicting the fact that G avoids P.

By the inductive hypothesis, it follows that:

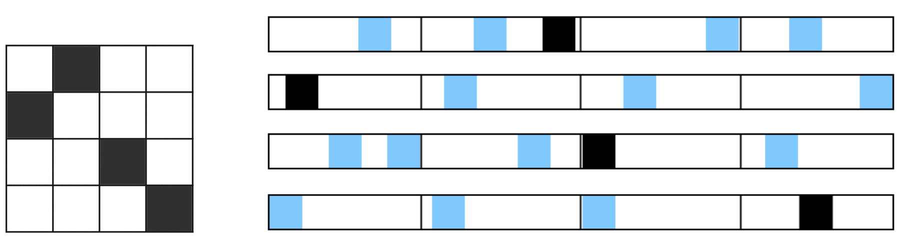
Fact 1. *There are at most $X_{\frac{n}{m}}$ black blocks.*

Each black block contains at most m^2 black squares, and each white block contains 0 black squares. Thus, we obtain the following recurrence:

$$X_n \leq X_{\frac{n}{m}} \cdot m^2$$

Unfortunately this recurrence is too weak. It will only prove $X_n \leq C \cdot n^2$.

The second key idea is to show that most black blocks have less than m^2 black squares. Indeed look at one row of blocks in G, call it R. Suppose that k different blocks in R all have a black square in the same k rows of squares. Then we can choose a black square from each of these rows, one in each block column, that has the same shape as P, contradiction! (The picture below and to the right shows an example of 4 rows picked out of 4 blocks, where each block has a black square in each row. Just like we did, you can find squares matching any 4×4 pattern.)



Now we're in business! Let's say a block is *tall* if it has at least k rows with a black square. How many tall blocks can there be altogether in R? As shown above, there are at most k tall blocks for each choice of k rows. R itself has m rows, so there are $\binom{m}{k}$ choices of k rows. Thus the total number of tall blocks in R is at most $\binom{m}{k} \cdot k$. Finally there are $\frac{n}{m}$ rows of blocks. Adding across everything, we get:

Fact 2. *There are at most $\frac{n}{m} \cdot \binom{m}{k} \cdot k$ tall blocks.*

Similarly, we'll call a block *wide* if it has at least k columns with a block square. Then, the same argument gives us:

Fact 3. *There are at most $\frac{n}{m} \cdot \binom{m}{k} \cdot k$ wide blocks.*

We can now get a better upper bound for X_n:

- There are at most $X_{\frac{n}{m}}$ black blocks that are neither tall nor wide. Each such block has at most $k-1$ rows with a black square and at most $k-1$ columns with a black square, and so the block itself has at most $(k-1)^2$ black squares.
- There are at most $\frac{n}{m} \cdot \binom{m}{k} \cdot k$ tall blocks, each containing at most m^2 black squares.
- There are at most $\frac{n}{m} \cdot \binom{m}{k} \cdot k$ wide blocks, each containing at most m^2 black squares.

Therefore, the total number of black squares is at most

$$X_{\frac{n}{m}} \cdot (k-1)^2 + 2 \cdot \frac{n}{m} \cdot \binom{m}{k} \cdot k \cdot m^2.$$

And now we're basically done! We are going to need our bound for $X_{\frac{n}{m}} \cdot (k-1)^2$ to be smaller than our bound for X_n, which means we are going to need to choose $m > (k-1)^2$. Fortunately, m can be whatever we want, and we will choose[2] to take $m = k^2$. This gives

$$X_n \leq X_{\frac{n}{k^2}} \cdot (k-1)^2 + 2nk^3 \cdot \binom{k^2}{k}.$$

Finally we will prove by induction that $X_n \leq 2k^4 \cdot \binom{k^2}{k} \cdot n$ for all n. This is obviously true $n=1$, and if the result holds for all values less than n, then we have

$$\begin{aligned} X_n &\leq 2k^4 \cdot \binom{k^2}{k} \cdot \frac{n}{k^2} \cdot (k-1)^2 + 2nk^3 \cdot \binom{k^2}{k} \\ &= 2k^4 \cdot \binom{k^2}{k} \cdot \left(\frac{(k-1)^2}{k^2} + \frac{1}{k} \right) \cdot n \\ &\leq 2k^4 \cdot \binom{k^2}{k} \cdot n, \end{aligned}$$

and now we really are done![3] This is more difficult than anything you would see on an Olympiad of course, but still very much the same kinds of ideas.

[2] Where exactly does this come from, and what about $2k^4 \cdot \binom{k^2}{k} \cdot n$ come from in the final part below? The idea is this: we are going to need to prove a bound of the form $Cn \leq C \cdot \frac{n}{m} \cdot (k-1)^2 + 2nmk \cdot \binom{k^2}{k}$. You can re-arrange this to figure out what C needs to be.

[3] We cheated in one way – we assume that n is a multiple of k^2. To get around this, we can use the fact that $X_{n+1} \leq X_n + 2n - 1$ and reduce first to a multiple of k^2 before doing the recurrence. It doesn't change anything really.

4 Problems

Here's a bunch of problems, organized as best I can in increasing order of difficulty. Some are easier than the examples, but there are also some really tough ones at the end!

1. Neal wants to buy exactly n dollars worth of snacks for the winter camp. He has an infinite number of two-dollar coins and an infinite number of five-dollar bills. Show that for every positive integer $n \geq 4$, he will be able to pay for the snacks without requiring any change.

2. Let $\mathbb{N} = \{1, 2, 3, \ldots\}$ be the set of positive integers. Find all functions f, defined on $\mathbb{N}$ and taking values in $\mathbb{N}$, such that $(n-1)^2 < f(n)f(f(n)) < n^2 + n$ for every positive integer n.

3. Let k be a positive integer. Prove that there exist integers x and y, neither of which is divisible by 3, such that $x^2 + 2y^2 = 3^k$.

4. Call an integer square-full if each of its prime factors occurs to a second power (at least). Prove that there are infinitely many pairs of consecutive square-fulls.

5. Let $a_1, a_2, \ldots, a_n$ be an increasing sequence of real numbers, and let $b_1, b_2, \ldots, b_n$ be another sequence that contains the same numbers but possibly in a different order. Suppose

$$a_1 + b_1 < a_2 + b_2 < \ldots < a_n + b_n.$$

   Prove that $b_i = a_i$ for all i.

6. For every finite set of positive integers A, define the collection $S(A)$ to be

$$S(A) = \{a + b \mid a, b \in A, a \neq b\}.$$

   For example, if $A = \{1, 2, 3, 4\}$, then $S(A) = \{3, 4, 5, 5, 6, 7\}$. Note that we are allowed to repeat elements in a collection.) Prove that there are infinitely many positive integers n for which we can find distinct sets A, B with $|A| = |B| = n$ and $S(A) = S(B)$.

7. Inside a right triangle a finite set of points is given. Prove that these points can be connected by a broken line such that the sum of the squares of the lengths in the broken line is less than or equal to the square of the length of the hypotenuse of the given triangle.

8. James and Matthew play a game that starts with two positive integers written on a board. They take turns, starting with James. On a player's turn, he subtracts the smaller number from the larger number one or more times. The first player to write a number that is 0 or less loses. For which pairs of integers does James have a winning strategy?

9. Let k be positive integer and m be odd number. Prove that there exists positive integer n such that $n^n - m$ is divisible by 2^k.

10. Let $n \geq k \geq 1$ be integers. Divide all the numbers with less than n digits into two classes, those whose digits add up to an even number and those whose digits add up to an odd number. Prove that the sum of the k-th powers of all numbers in one class is equal to the sum of the k-th powers of all numbers in the other class.

11. I'm playing the colour-country game against Dorette. We take turns; on my turn, I draw in a country. On Dorette's turn, she chooses any colour for the country, but he must make sure that no adjacent countries share the same colour. Is it possible for me to force Dorette to use more than 9000 colors?

12. Using the notation from Section 3, fix a pattern P. Let Y_n denote the number of $n \times n$ grids that avoid P. Prove that there exists a real number C such that $Y_n \leq C^n$ for all n. (Use the theorem – do not try to prove this directly!!)

13. Show that the set of positive integers that cannot be represented as a sum of distinct perfect squares is finite.

14. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_m$ be positive integers such that $a_1 + a_2 + \ldots + a_n = b_1 + b_2 + \ldots + b_m < nm$. If $n, m \geq 2$, prove that it is possible to delete some, but not all, of the a_i's and b_j's so that the sum is still equal.

15. (a) Consider a polygon with some non-intersecting diagonals drawn in. Prove it is possible to colour the vertices red, blue, or green, such that two vertices of the same colour are never connected by an edge or a diagonal.
    (b) An art gallery is shaped like a polygon with n sides. Prove that we can place at most $\frac{n}{3}$ guards inside so that every other point inside the gallery is visible to at least one guard.

16. Each vertex of a finite graph can be coloured either black or white. Initially all vertices are black. We are allowed to pick a vertex P and change the colour of P and all of its neighbours. Is it always possible to change the colour of every vertex from black to white by a sequence of operations of this type?

17. Determine all functions f from the set of positive integers to the set of positive integers such that, for all positive integers a and b, there exists a non-degenerate triangle with sides of lengths
$$a, f(b), \text{ and } f(b + f(a) - 1).$$
(A triangle is non-degenerate if its vertices are not colinear.)

18. Let $a_1, a_2, \ldots, a_n$ be a sequence of distinct positive real numbers. For the next n days, there will be a snowstorm each day covering all positions on the x-axis with value ≥ 0 *or* all positions with value ≤ 0. Sarah starts at position 0 and knows where each snowstorm will be. Before each snowstorm, Sarah will jump some distance to one side to make sure she avoids it (even if she was already avoiding it). Prove that no matter where the snowstorms are going to be, it is always possible for Sarah to do this using exactly one jump of length i for each $i \in \{1, 2, \ldots, n\}$.

19. Back in the 2000 Winter Camp, we invited 512 students and assigned them to 256 rooms with 2 students per room. We also had 9 tests, and it turned out that no two students aced the exact same set of tests. Prove that it was possible to arrange all 512 students into a circle such that (a) all roommates were next to each other, and (b) if two non-roommates were next to each other, then one of them aced all the tests that the other student did, plus exactly one additional test.

20. Let $\mathbb{Z}_{>0}$ denote the set of positive integers. For any positive integer k, a function $f : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ is called k-good if $\gcd(f(m)+n, f(n)+m) \leq k$ for all $m \neq n$. Find all k such that there exists a k-good function.

21. An ordered pair (x, y) of integers is a primitive point if the greatest common divisor of x and y is 1. Given a finite set S of primitive points, prove that there exist a positive integer n and integers $a_0, a_1, \ldots, a_n$ such that, for each (x, y) in S, we have:

$$a_0x^n + a_1x^{n-1}y + a_2x^{n-2}y^2 + \cdots + a_{n-1}xy^{n-1} + a_ny^n = 1.$$

22. Let $a_1, a_2, \ldots, a_n$ be distinct positive integers and let M be a set of $n-1$ positive integers not containing $s = a_1 + a_2 + \ldots + a_n$. A grasshopper is to jump along the real axis, starting at the point 0 and making n jumps to the right with lengths $a_1, a_2, \ldots, a_n$ in some order. Prove that the order can be chosen in such a way that the grasshopper never lands on any point in M.

5 Hints

1. If it's possible to do $(k, k+1, k+2, k+3, k+4)$, prove it's also possible to do $k+5$.
2. (CMO 2015 #1) Suppose $f(n) = n$ for $n \in \{1, 2, \ldots, N-1\}$; then we want to show $f(N) = N$. If $f(N) < N$, what is $f(N)f(f(N))$? Conversely if $f(N) = M > N$, what is $f(M)f(f(M))$?
3. (Bay Area Math Olympiad 2015 #3) You can try solving for small k and looking for patterns. Alternatively, suppose $x^2 + 2y^2 = 3^k$ and consider $(Ax + By)^2 + 2(Cx + Dy)^2$.
4. Suppose a and $a+1$ are square-full; you're going to want to use that pair to generate a bigger pair. Think about $a(a+1)$. Does that give you any ideas?
5. Suppose $a_1 \notin \{b_1, b_2, \ldots, b_k\}$. Prove that $a_1 + b_{k+1} < a_k + b_k$ and hence $a_1 \neq b_{k+1}$ also.
6. Suppose A, B have the desired property. Try to use A, B to form two sets that are twice as large. Note that n *must* be a power of 2!
7. How can you divide a right triangle into two smaller right triangles? What does that suggest?
8. Prove that James has a winning strategy if and only if $\max(n, m) \geq \sqrt{2} \cdot \min(n, m)$.
9. Suppose $2^k | n^n - m$. If $n^n \not\equiv m \mod 2^{k+1}$, then what is it?
10. This requires some good old fashioned algebra. Let $s_{n,k}$ be the sum of the k^{th} powers of all numbers with less than n digits, but negating all numbers with odd digit sum. Try separating out the last digit and reducing to $s_{n-1,k}$.
11. You need to strengthen the inductive hypothesis: prove I can force Dorette to use n different colours for countries that are on a boundary somewhere.
12. Divide the grid into blocks of size 2×2 and try to bound Y_n by $Y_{n/2}$. By the way, this is the real Stanley-Wilf conjecture, not the theorem in Section 3.
13. (IMO Shortlist 2000 N6) The main challenge here is there is no good base case. One approach is to try to reduce to a smaller n but one that is still big enough. I think it's cleaner to work in modular residue classes though. Every number that is 5 mod 16 can be written as a sum of unique squares, and this is straightforward. Do you see how to finish from there?
14. Let a_n and b_m be the largest numbers in their respective sequences, and assume without loss of generality that $a_n > b_m$. Try deleting them both and adding $a_n - b_m$ to the a list instead.
15. This is a famous problem, called the Art Gallery problem. (a) is a pretty straightforward induction. For (b), you need to prove every n-sided polygon can be divided into $n-2$ triangles by adding diagonals. You can do this by induction too – if ABC are consecutive vertices, you can either connect AC or connect B to a vertex inside triangle ABC. After that, you want to put the guards at some of the vertices; use the colouring to decide how.
16. (CMO 2010 #4, but the result was known before that) Using an inductive hypothesis, you can flip the colour of every vertex but one, and you do not know what happens to the final vertex. If none of these options solve the problem outright, use this fact twice to prove you can flip any two vertices, and go from there.

17. (IMO 2009 #5) After proving $f(1) = 1$, let $x_n = 1 + n\big(f(2) - 1\big)$. Prove by induction that $f(x_n) = n + 1$.

18. (IOI 2002) Strengthen the inductive hypothesis. Let $a_1 < a_2 < a_3 < \ldots$. Prove that it is possible for Sarah to accomplish her task if she makes $a_1, a_3, a_5, \ldots$ go one direction and $a_2, a_4, a_6, \ldots$ go the other direction.

19. (Russia Grade 11 2010 #8) Consider all students who aced test #1. If you can pair them up in some way without creating any cycles, then you can reduce the problem from 2^n students to 2^{n-1} students, considering only those who did not ace test #1. You can use the inductive hypothesis to find this initial pairing too.

20. (IMO Shortlist 2015 N7) James wrote this problem! Such a function exists for $k \geq 2$. I think the most natural approach is to define f inductively. Suppose we have defined $f(1), f(2), \ldots, f(n-1)$ and are looking to define $f(n)$. Let P be the set of primes (or 4) that divide $f(m) + n$ for $m \in \{1, 2, \ldots, n-1\}$. Then it suffices to choose $f(n)$ such that nothing in P divides $f(n) + m$. A counterexample would have $p | f(n) + n + f(m) + m$, so this suggests choosing $f(n) = C - n$ for some C for each prime p in P. You can do this by the Chinese remainder theorem. You just have to make sure $f(m) + m$ hasn't already used by all the possible spots.

21. (IMO 2017 #6) Start with a solution $P(x)$ for $(p_1, q_1), (p_2, q_2), \ldots, (p_n, p_n)$, and think about how you can easily generate more solutions. Consider $P(x)^N - x^{N-n} \cdot \prod(q_i x - p_i y)$. Intuitively it helps a lot if you assume $(p_{n+1}, q_{n+1}) = (1, 0)$ and you actually can do this – it comes down to a change of variables.

22. (IMO 2009 #6) Let a_n be the largest jump. Inductively start with a solution for $n - 1$ that ignores a_n, and then add a_n at the end. The only way this can fail is if the last jump before a_n hits a bad spot. But you could swap that last jump with a_n, and *that* only fails if there's two bad spots near the end. But that helps set up induction another way.