Chapter 2

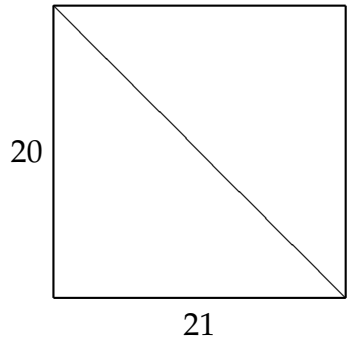
Induction I

2.1 A Warmup Puzzle

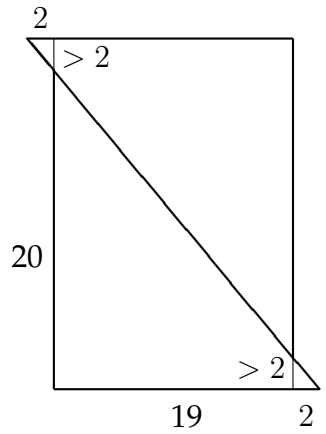
In principle, a proof should establish the truth of a proposition with absolute certainty. In practice, however, many purported proofs contain errors: overlooked cases, logical slips, and even algebra mistakes. But in a well-written proof, even if there is a bug, one should at least be able to pinpoint a specific statement that does not logically follow. See if you can find the first error in the following argument.

False Theorem 11. $420 > 422$

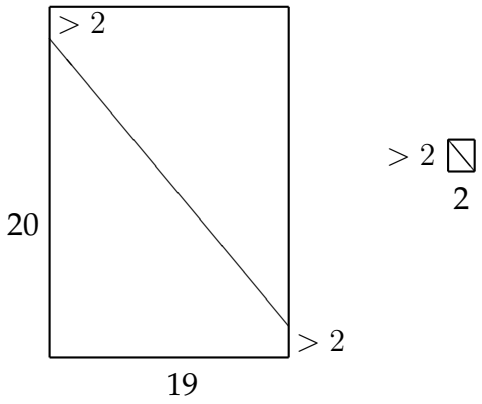
Proof. We will demonstrate this fact geometrically. We begin with a 20×21 rectangle, which has area 420:



Now we cut along the diagonal as indicated above and slide the upper piece parallel to the cut until it has moved exactly 2 units leftward. This leaves a couple stray corners, which are 2 units wide and just over 2 units high.



Finally, we snip off the two corners and place them together to form an additional small rectangle:



Now we have two rectangles, a large one with area just over $(20 + 2) \times 19 = 418$ and a small one with area just over $2 \times 2 = 4$. Thus, the total area of the resulting figure is a bit over $418 + 4 = 422$. By conservation of area, 420 is equal to just a little bit more than 422. □

Where is the error?

2.2 Induction

A professor brings to class a bottomless bag of assorted miniature candy bars. She offers to share in accordance with two rules. First, she numbers the students 0, 1, 2, 3, and so forth for convenient reference. Now here are the two rules:

1. Student 0 gets candy.

2. For all $n \in \mathbb{N}$, if student n gets candy, then student $n + 1$ also gets candy.

You can think of the second rule as a compact way of writing a whole sequence of statements, one for each natural value of n:

- If student 0 gets candy, then student 1 also gets candy.
- If student 1 gets candy, then student 2 also gets candy.
- If student 2 gets candy, then student 3 also gets candy, and so forth.

Now suppose you are student 17. By these rules, are you entitled to a miniature candy bar? Well, student 0 gets candy by the first rule. Therefore, by the second rule, student 1 also gets candy, which means student 2 gets candy as well, which means student 3 get candy, and so on. So the professor's two rules actually guarantee candy for *every* student, no matter how large the class. You win!

This reasoning generalizes to a principle called ***induction***:

Principle of Induction. Let $P(n)$ be a predicate. If

- $P(0)$ is true, and
- for all $n \in \mathbb{N}$, $P(n)$ implies $P(n + 1)$,

then $P(n)$ is true for all $n \in \mathbb{N}$.

Here's the correspondence between the induction principle and sharing candy bars. Suppose that $P(n)$ is the predicate, "student n gets candy". Then the professor's first rule asserts that $P(0)$ is true, and her second rule is that for all $n \in \mathbb{N}$, $P(n)$ implies $P(n + 1)$. Given these facts, the induction principle says that $P(n)$ is true for all $n \in \mathbb{N}$. In other words, everyone gets candy.

The intuitive justification for the general induction principle is the same as for everyone getting a candy bar under the professor's two rules. Mathematicians find this intuition so compelling that induction is always either taken as an axiom or else proved from more primitive axioms, which are themselves specifically designed so that induction *is* provable. In any case, the induction principle is a core truth of mathematics.

2.3 Using Induction

Induction is by far the most important proof technique in computer science. Generally, induction is used to prove that some statement holds for all natural values of a variable. For example, here is a classic formula:

Theorem 12. *For all $n \in \mathbb{N}$:*

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

The left side of the equation represents the sum of all the numbers from 1 to n. You're supposed to guess the pattern and mentally replace the ... with the other terms. We could eliminate the need for guessing by rewriting the left side with ***summation notation***:

$$\sum_{i=1}^{n} i \qquad \text{or} \qquad \sum_{1 \le i \le n} i \qquad \text{or} \qquad \sum_{i \in \{1,\ldots,n\}} i$$

Each of these expressions denotes the sum of all values taken on by the expression to the right of the sigma as the variable i ranges from 1 to n. The meaning of the sum in Theorem 12 is not so obvious in a couple special cases:

- If $n = 1$, then there is only one term in the summation, and so $1+2+3+\ldots+n = 1$. Don't be misled by the appearance of 2 and 3 and the suggestion that 1 and n are distinct terms!
- If $n \le 0$, then there are no terms at all in the summation, and so $1+2+3+\ldots+n = 0$.

The ... notation is convenient, but watch out for these special cases where the notation is misleading!

Now let's use the induction principle to prove Theorem 12. Suppose that we define predicate $P(n)$ to be "$1 + 2 + 3 + \ldots + n = n(n+1)/2$". Recast in terms of this predicate, the theorem claims that $P(n)$ is true for all $n \in \mathbb{N}$. This is great, because the induction principle lets us reach precisely that conclusion, provided we establish two simpler facts:

- $P(0)$ is true.
- For all $n \in \mathbb{N}$, $P(n)$ implies $P(n+1)$.

So now our job is reduced to proving these two statements. The first is true because $P(0)$ asserts that a sum of zero terms is equal to $0(0+1)/2 = 0$.

The second statement is more complicated. But remember the basic plan for proving the validity of any implication: *assume* the statement on the left and then *prove* the statement on the right. In this case, we assume $P(n)$:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

in order to prove $P(n+1)$:

$$1 + 2 + 3 + \ldots + n + (n+1) = \frac{(n+1)(n+2)}{2}$$

These two equations are quite similar; in fact, adding $(n+1)$ to both sides of the first equation and simplifying the right side gives the second equation:

$$\begin{aligned} 1+2+3+\ldots+n+(n+1) &= \frac{n(n+1)}{2}+(n+1) \\ &= \frac{(n+2)(n+1)}{2} \end{aligned}$$

Thus, if $P(n)$ is true, then so is $P(n+1)$. This argument is valid for every natural number n, so this establishes the second fact required by the induction principle. In effect, we've just proved that $P(0)$ implies $P(1)$, $P(1)$ implies $P(2)$, $P(2)$ implies $P(3)$, etc. all in one fell swoop.

With these two facts in hand, the induction principle says that the predicate $P(n)$ is true for all natural n. And so the theorem is proved!

A Template for Induction Proofs

The proof of Theorem 12 was relatively simple, but even the most complicated induction proof follows exactly the same template. There are five components:

1. **State that the proof uses induction.** This immediately conveys the overall structure of the proof, which helps the reader understand your argument.

2. **Define an appropriate predicate $P(n)$.** The eventual conclusion of the induction argument will be that $P(n)$ is true for all natural n. Thus, you should define the predicate $P(n)$ so that your theorem is equivalent to (or follows from) this conclusion. Often the predicate can be lifted straight from the claim, as in the example above. The predicate $P(n)$ is called the "induction hypothesis".

3. **Prove that $P(0)$ is true.** This is usually easy, as in the example above. This part of the proof is called the "base case" or "basis step".

4. **Prove that $P(n)$ implies $P(n+1)$ for every natural number n.** This is called the "inductive step" or "induction step". The basic plan is always the same: assume that $P(n)$ is true and then use this assumption to prove that $P(n+1)$ is true. These two statements should be fairly similar, but bridging the gap may require some ingenuity. Whatever argument you give must be valid for every natural number n, since the goal is to prove the implications $P(0) \to P(1)$, $P(1) \to P(2)$, $P(2) \to P(3)$, etc. all at once.

5. **Invoke induction.** Given these facts, the induction principle allows you to conclude that $P(n)$ is true for all natural n. This is the logical capstone to the whole argument, but many writers leave this step implicit.

Explicitly labeling the *base case* and *inductive step* may make your proofs more clear.

A Clean Writeup

The proof of Theorem 12 given above is perfectly valid; however, it contains a lot of extraneous explanation that you won't usually see in induction proofs. The writeup below is closer to what you might see in print and should be prepared to produce yourself.

Proof. We use induction. Let $P(n)$ be the predicate:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

Base case: $P(0)$ is true, because both sides of the equation are zero.

Inductive step: Assume that $P(n)$ is true, where n is any natural number. Then $P(n+1)$ is also true, because:

$$\begin{aligned} 1 + 2 + 3 + \ldots + n + (n+1) &= \frac{n(n+1)}{2} + (n+1) \\ &= \frac{(n+1)(n+2)}{2} \end{aligned}$$

The first step uses the assumption $P(n)$, and the second follows by simplification.

Therefore, $P(n)$ is true for all natural n by induction, and the theorem is proved. □

Induction was helpful for *proving the correctness* of this summation formula, but not helpful for *discovering* the formula in the first place. There are some tricks for finding such formulas, which we'll show you in a few weeks.

2.4 A Divisibility Theorem

An integer a ***divides*** an integer b if b is a multiple of a. This is denoted $a \mid b$. For example, $3 \mid (5^3 - 5)$, since $5^3 - 5 = 120$ is a multiple of 3. More generally, we have the following theorem:

Theorem 13. $\forall\, n \in \mathbb{N} \quad 3 \mid (n^3 - n)$

Let's try to prove this with induction. The first challenge is always selecting the right induction hypothesis, $P(n)$. Your first instinct should be to lift the induction hypothesis directly from the claim. Thus, in this case, we should first try letting $P(n)$ be the predicate "$3 \mid (n^3 - n)$". (This doesn't always work out – as we'll see in a later example – but it does work in this case.)

Now we must address the base case by proving that $P(0)$ is true. As is often the case, this is easy: $3 \mid (0^3 - 0)$, since 0 is a multiple of 3. (Specifically, $3 \cdot 0 = 0$.)

Our next task, the inductive step, is typically the most difficult part of an induction proof. We must show that $P(n)$ implies $P(n+1)$. Thus, as usual when proving an implication, we *assume* $P(n)$ in order to *prove* $P(n+1)$. Writing out what these two expressions actually mean is often helpful. In this case, we assume $P(n)$:

$$3 \mid (n^3 - n)$$

in order to prove $P(n+1)$:

$$3 \mid ((n+1)^3 - (n+1))$$

These two statements look somewhat similar, but how can we use the first to prove the second? For lack of any other ideas, let's multiply out the expression in the second statement:

$$\begin{aligned} 3 \mid ((n+1)^3 - (n+1)) &\Leftrightarrow 3 \mid (n^3 + 3n^2 + 3n + 1 - n - 1) \\ &\Leftrightarrow 3 \mid (n^3 + 3n^2 + 2n) \end{aligned}$$

Aha! Notice that the last expression is equal to $n^3 - n$ plus $3n^2 + 3n$. Since $3n^2 + 3n$ is a multiple of 3 and $n^3 - n$ is a multiple of 3 by assumption, their sum must also be a multiple of 3. Therefore, $((n+1)^3 - (n+1)$ must also be a multiple of 3.

Playing around with $P(n)$ and $P(n+1)$ in this way, trying to understand how the two are related, is pretty typical when one is searching for an induction argument. However, what we've done so far is only scratchwork. What remains is to organize our reasoning into a clear proof.

Proof. We use induction. Let $P(n)$ be the proposition that $3 \mid (n^3 - n)$.

Base case: $P(0)$ is true because $3 \mid 0^3 - 0$.

Inductive step: Assume that $P(n)$ is true, where n is any natural number. Then:

$$\begin{aligned} 3 \mid (n^3 - n) &\Rightarrow 3 \mid (n^3 - n) + 3(n^2 + n) \\ &\Rightarrow 3 \mid n^3 + 3n^2 + 3n + 1 - n - 1 \\ &\Rightarrow 3 \mid (n+1)^3 - (n+1) \end{aligned}$$

The first implication relies on the fact that $3(n^2 + n)$ is divisible by 3. The remaining implications involve only rewriting the expression on the right. The last statement is $P(n+1)$, so we've proved that $P(n)$ implies $P(n+1)$ for all $n \in \mathbb{N}$.

By the principle of induction, $P(n)$ is true for all $n \in \mathbb{N}$, which proves the claim. □

This proof would look quite mysterious to anyone not privy to the scratchwork we did beforehand. In particular, one might ask how we had the foresight to introduce the magic term $3(n^2 + n)$. Of course, this was not foresight at all; we just worked backward initially!

2.5 A Faulty Induction Proof

Sometimes we want to prove that a statement is true for, say, all integers $n \geq 1$ rather than all integers $n \geq 0$. In this circumstance, we can use a slight variation on induction: prove $P(1)$ in the base case and then prove that $P(n)$ implies $P(n+1)$ for all $n \geq 1$ in the inductive step. This is a perfectly valid variant of induction and is *not* the problem with the proof below.

False Theorem 14. *All horses are the same color.*

Proof. The proof is by induction. Let $P(n)$ be the proposition that in every set of n horses, all are the same color.

Base case: $P(1)$ is true, because all horses in a set of 1 must be the same color.

Inductive step: Assume that $P(n)$ is true, where n is a positive integer; that is, assume that in every set of n horses, all are the same color. Now consider a set of $n+1$ horses:

$$h_1,\ h_2,\ \ldots,\ h_n,\ h_{n+1}$$

By our assumption, the first n horses are the same color:

$$\underbrace{h_1,\ h_2,\ \ldots,\ h_n,}_{\text{same color}}\ h_{n+1}$$

Also by our assumption, the last n horses are the same color:

$$h_1,\ \underbrace{h_2,\ \ldots,\ h_n,\ h_{n+1}}_{\text{same color}}$$

Therefore, horses $h_1, h_2, \ldots, h_{n+1}$ must all be the same color, and so $P(n+1)$ is true. Thus, $P(n)$ implies $P(n+1)$.

By the principle of induction, $P(n)$ is true for all $n \geq 1$. The theorem is a special case where n is equal to the number of horses in the world. □

We've proved something false! Is math broken? Should we all become poets?

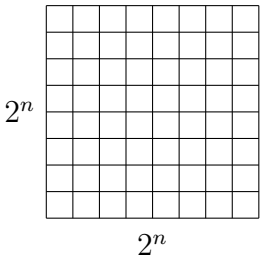
The error in this argument is in the sentence that begins, "Therefore, horses h_1, h_2, …, h_n, h_{n+1} must all be the same color." The "…" notation creates the impression that the sets $h_1, h_2, \ldots, h_n$ and $h_2, \ldots, h_n, h_{n+1}$ overlap. However, this is not true when $n = 1$. In that case, the first set is just h_1 and the second is h_2, and these do not overlap at all!

This mistake knocks a critical link out of our induction argument. We proved $P(1)$ and we proved $P(2) \Rightarrow P(3)$, $P(3) \Rightarrow P(4)$, etc. But we failed to prove $P(1) \Rightarrow P(2)$, and so everything falls apart: we can not conclude that $P(3)$, $P(4)$, etc. are true. And, of course, these propositions are all false; there are horses of a different color.

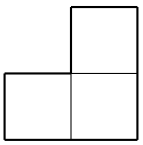
2.6 Courtyard Tiling

Induction served purely as a proof technique in the preceding examples. But induction sometimes can serve as a more general reasoning tool.

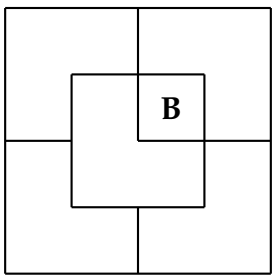
MIT recently constructed a new computer science building. As the project went further and further over budget, there were some radical fundraising ideas. One plan was to install a big courtyard with dimensions $2^n \times 2^n$:



One of the central squares would be occupied by a statue of a wealthy potential donor. Let's call him "Bill". (In the special case $n = 0$, the whole courtyard consists of a single central square; otherwise, there are four central squares.) A complication was that the building's unconventional architect, Frank Gehry, insisted that only special L-shaped tiles be used:



A courtyard meeting these constraints exsists, at least for $n = 2$:



For larger values of n, is there a way to tile a $2^n \times 2^n$ courtyard with L-shaped tiles and a statue in the center? Let's try to prove that this is so.

Theorem 15. *For all $n \geq 0$ there exists a tiling of a $2^n \times 2^n$ courtyard with Bill in a central square.*

Proof. (doomed attempt) The proof is by induction. Let $P(n)$ be the proposition that there exists a tiling of a $2^n \times 2^n$ courtyard with Bill in the center.

Base case: $P(0)$ is true because Bill fills the whole courtyard.

Inductive step: Assume that there is a tiling of a $2^n \times 2^n$ courtyard with Bill in the center for some $n \geq 0$. We must prove that there is a way to tile a $2^{n+1} \times 2^{n+1}$ courtyard with Bill in the center... □

Now we're in trouble! The ability to tile a smaller courtyard with Bill in the center does not help tile a larger courtyard with Bill in the center. We can not bridge the gap between $P(n)$ and $P(n+1)$. The usual recipe for finding an inductive proof will not work!

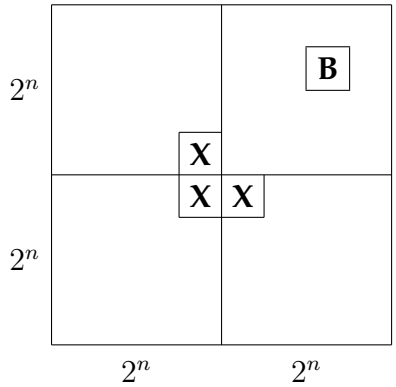
When this happens, your first fallback should be to look for a *stronger* induction hypothesis; that is, one which implies your previous hypothesis. For example, we could make $P(n)$ the proposition that for *every* location of Bill in a $2^n \times 2^n$ courtyard, there exists a tiling of the remainder.

This advice may sound bizzare: "If you can't prove something, try to prove something more grand!" But for induction arguments, this makes sense. In the inductive step, where you have to prove $P(n) \Rightarrow P(n+1)$, you're in better shape because you can *assume* $P(n)$, which is now a more general, more useful statement. Let's see how this plays out in the case of courtyard tiling.

Proof. (successful attempt) The proof is by induction. Let $P(n)$ be the proposition that for every location of Bill in a $2^n \times 2^n$ courtyard, there exists a tiling of the remainder.

Base case: $P(0)$ is true because Bill fills the whole courtyard.

Inductive step: Asume that $P(n)$ is true for some $n \geq 0$; that is, for every location of Bill in a $2^n \times 2^n$ courtyard, there exists a tiling of the remainder. Divide the $2^{n+1} \times 2^{n+1}$ courtyard into four quadrants, each $2^n \times 2^n$. One quadrant contains Bill (**B** in the diagram below). Place a temporary Bill (**X** in the diagram) in each of the three central squares lying outside this quadrant:



Now we can tile each of the four quadrants by the induction assumption. Replacing the three temporary Bills with a single L-shaped tile completes the job. This proves that $P(n)$ implies $P(n+1)$ for all $n \geq 0$. The theorem follows as a special case. □

This proof has two nice properties. First, not only does the argument guarantee that a tiling exists, but also it gives an algorithm for finding such a tiling. Second, we have a stronger result: if Bill wanted a statue on the edge of the courtyard, away from the pigeons, we could accommodate him!

Strengthening the induction hypothesis is often a good move when an induction proof won't go through. But keep in mind that the stronger assertion must actually be *true*; otherwise, there isn't much hope of constructing a valid proof! Sometimes finding just the right induction hypothesis requires trial, error, and insight. For example, mathematicians spent almost twenty years trying to prove or disprove the conjecture that "Every planar graph is 5-choosable"[1]. Then, in 1994, Carsten Thomassen gave an induction proof simple enough to explain on a napkin. The key turned out to be finding an extremely clever induction hypothesis; with that in hand, completing the argument is easy!

2.7 Another Faulty Proof

False Theorem 16. *I can lift all the sand on the beach.*

Proof. The proof is by induction. Let $P(n)$ be the predicate, "I can lift n grains of sand." The base case $P(1)$ is true because I can certainly lift one grain of sand. In the inductive step, assume that I can lift n grains of sand to prove that I can lift $n+1$ grains of sand. If I can lift n grains of sand, then surely I can lift $n+1$; one grain of sand will not make any difference. Therefore $P(n) \Rightarrow P(n+1)$. By induction, $P(n)$ is true for all $n \geq 1$. The theorem is the special case where n is equal to the number of grains of sand on the beach. □

The flaw here is in the bogus assertion that I can lift $n+1$ grains of sand because I can lift n grains. It is hard to say for exactly which n this is false, but certainly there is some value!

There is a field of mathematics called "fuzzy logic" in which truth is not a 0/1 thing, but is rather represented by a real value between 0 and 1. There is an analogue of induction in which the truth value decreases a bit with each implication. That might better model the situation here, since my lifts would probably gradually grow more and more sorry-looking as n approached my maximum. We will not be using fuzzy logic in this class, however. At least not intentionally.

[1] 5-choosability is a slight generalization of 5-colorability. Although every planar graph is 4-colorable and therefore 5-colorable, not every planar graph is 4-choosable. If this all sounds like nonsense, don't panic. We'll discuss graphs, planarity, and coloring in two weeks.

Chapter 3

Induction II

3.1 Good Proofs and Bad Proofs

In a purely technical sense, a mathematical proof is verification of a proposition by a chain of logical deductions from a base set of axioms. But the *purpose* of a proof is to provide readers with compelling evidence for the truth of an assertion. To serve this purpose effectively, more is required of a proof than just logical correctness: a good proof must also be clear. These goals are complimentary; a well-written proof is more likely to be a correct proof, since mistakes are harder to hide. Here are some tips on writing good proofs:

State your game plan. A good proof begins by explaining the general line of reasoning, e.g. "We use induction" or "We argue by contradiction". This creates a rough mental picture into which the reader can fit the subsequent details.

Keep a linear flow. We sometimes see proofs that are like mathematical mosaics, with juicy tidbits of reasoning sprinkled judiciously across the page. This is not good. The steps of your argument should follow one another in a clear, sequential order.

Explain your reasoning. Many students initially write proofs the way they compute integrals. The result is a long sequence of expressions without explantion. This is bad. A good proof usually looks like an essay with some equations thrown in. Use complete sentences.

Avoid excessive symbolism. Your reader is probably good at understanding words, but much less skilled at reading arcane mathematical symbols. So use words where you reasonably can.

Simplify. Long, complicated proofs take the reader more time and effort to understand and can more easily conceal errors. So a proof with fewer logical steps is a better proof.

Introduce notation thoughtfully. Sometimes an argument can be greatly simplified by introducing a variable, devising a special notation, or defining a new term. But do this sparingly, since you're requiring the reader to remember all this new stuff. And remember to actually *define* the meanings of new variables, terms, or notations; don't just start using them.

Structure long proofs. Long program are usually broken into a heirarchy of smaller procedures. Long proofs are much the same. Facts needed in your proof that are easily stated, but not readily proved are best pulled out and proved in a preliminary lemmas. Also, if you are repeating essentially the same argument over and over, try to capture that argument in a general lemma and then repeatedly cite that instead.

Don't bully. Words such as "clearly" and "obviously" serve no logical function. Rather, they almost always signal an attempt to bully the reader into accepting something which the author is having trouble justifying rigorously. Don't use these words in your own proofs and go on the alert whenever you read one.

Finish. At some point in a proof, you'll have established all the essential facts you need. Resist the temptation to quit and leave the reader to draw the right conclusions. Instead, tie everything together yourself and explain why the original claim follows.

The analogy between good proofs and good programs extends beyond structure. The same rigorous thinking needed for proofs is essential in the design of critical computer system. When algorithms and protocols only "mostly work" due to reliance on handwaving arguments, the results can range from problematic to catastrophic. An early example was the Therac 25, a machine that provided radiation therapy to cancer victims, but occasionally killed them with massive overdoses due to a software race condition. More recently, a buggy electronic voting system credited presidential candidate Al Gore with *negative* 16,022 votes in one county. In August 2004, a single faulty command to a computer system used by United and American Airlines grounded the entire fleet of both companies— and all their passengers.

It is a certainty that we'll all one day be at the mercy of critical computer systems designed by you and your classmates. So we really hope that you'll develop the ability to formulate rock-solid logical arguments that a system actually does what you think it does.

3.2 A Puzzle

Here is a puzzle. There are 15 numbered tiles and a blank square arranged in a 4×4 grid. Any numbered tile adjacent to the blank square can be slid into the blank. For example, a sequence of two moves is illustrated below:

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

→

1	2	3	4
5	6	7	8
9	10	11	12
13	15		**14**

→

1	2	3	4
5	6	7	8
9	10		12
13	15	**11**	14

In the leftmost configuration shown above, the 14 and 15 tiles are out of order. Can you find a sequence of moves that puts these two numbers in the correct order, but returns every other tile to its original position? Some experimentation suggests that the answer is probably "no", so let's try to prove that.

We're going to take an approach that is frequently used in the analysis of software and systems. We'll look for an ***invariant***, a property of the puzzle that is always maintained, no matter how you move the tiles around. If we can then show that putting the 14 and 15 tiles in the correct order would violate the invariant, then we can conclude that this is impossible.

Let's see how this game plan plays out. Here is the theorem we're trying to prove:

Theorem 17. *No sequence of moves transforms the board below on the left into the board below on the right.*

1	2	3	4
5	6	7	8
9	10	11	12
13	**15**	**14**	

1	2	3	4
5	6	7	8
9	10	11	12
13	**14**	**15**	

We'll build up a sequence of observations, stated as lemmas. Once we achieve a critical mass, we'll assemble these observations into a complete proof.

Define a ***row move*** as a move in which a tile slides horizontally and a ***column move*** as one in which a tile slides vertically. Assume that tiles are read top-to-bottom and left-to-right like English text; so when we say two tiles are "out of order", we mean that the larger number precedes the smaller number in this order.

Our difficulty is that one pair of tiles (the 14 and 15) is out of order initially. An immediate observation is that row moves alone are of little value in addressing this problem:

Lemma 18. *A row move does not change the order of the tiles.*

Usually a lemma requires a proof. However, in this case, there are probably no more-compelling observations that we could use as the basis for such an argument. So we'll let this statement stand on its own and turn to column moves.

Lemma 19. *A column move changes the relative order of exactly 3 pairs of tiles.*

For example, the column move shown below changes the relative order of the pairs (j, g), (j, h), and (j, i).

a	b	c	d
e	f		g
h	i	**j**	k
l	m	n	o

$\rightarrow$

a	b	c	d
e	f	**j**	g
h	i		k
l	m	n	o

Proof. Sliding a tile down moves it after the next three tiles in the order. Sliding a tile up moves it before the previous three tiles in the order. Either way, the relative order changes between the moved tile and each of the three it crosses. □

These observations suggest that there are limitations on how tiles can be swapped. Some such limitation may lead to the invariant we need. In order to reason about swaps more precisely, let's define a term: a pair of tiles i and j is ***inverted*** if $i < j$, but i appears after j in the puzzle. For example, in the puzzle below, there are four inversions: (12, 11), (13, 11), (15, 11), and (15, 14).

1	2	3	4
5	6	7	8
9	10		12
13	15	11	14

Let's work out the effects of row and column moves in terms of inversions.

Lemma 20. *A row move never changes the parity of the number of inversions. A column move always changes the parity of the number of inversions.*

The "parity" of a number refers to whether the number is even or odd. For example, 7 and -5 have odd parity, and 18 and 0 have even parity.

Proof. By Lemma 18, a row move does not change the order of the tiles; thus, in particular, a row move does not change the number of inversions.

By Lemma 19, a column move changes the relative order of exactly 3 pairs of tiles. Thus, an inverted pair becomes uninverted and vice versa. Thus, one exchange flips the total number of inversions to the opposite parity, a second exhange flips it back to the original parity, and a third exchange flips it to the opposite parity again. □

This lemma implies that we must make an *odd* number of column moves in order to exchange just one pair of tiles (14 and 15, say). But this is problematic, because each column move also knocks the blank square up or down one row. So after an *odd* number of column moves, the blank can not possibly be back in the last row, where it belongs! Now we can bundle up all these observations and state an invariant, a property of the puzzle that never changes, no matter how you slide the tiles around.

Lemma 21. *In every configuration reachable from the position shown below, the parity of the number of inversions is different from the parity of the row containing the blank square.*

row 1	1	2	3	4
row 2	5	6	7	8
row 3	9	10	11	12
row 4	13	15	14	

Proof. We use induction. Let $P(n)$ be the proposition that after n moves, the parity of the number of inversions is different from the parity of the row containing the blank square.

Base case: After zero moves, exactly one pair of tiles is inverted (14 and 15), which is an odd number. And the blank square is in row 4, which is an even number. Therefore, $P(0)$ is true.

Inductive step: Now we must prove that $P(n)$ implies $P(n+1)$ for all $n \geq 0$. So assume that $P(n)$ is true; that is, after n moves the parity of the number of inversions is different from the parity of the row containing the blank square. There are two cases:

1. Suppose move $n+1$ is a row move. Then the parity of the total number of inversions does not change by Lemma 20. The parity of the row containing the blank square does not change either, since the blank remains in the same row. Therefore, these two parities are different after $n+1$ moves as well, so $P(n+1)$ is true.

2. Suppose move $n+1$ is a column move. Then the parity of the total number of inversions changes by Lemma 20. However, the parity of the row containing the blank square also changes, since the blank moves up or down one row. Thus, the parities remain different after $n+1$ moves, and so $P(n+1)$ is again true.

Thus, $P(n)$ implies $P(n+1)$ for all $n \geq 0$.

By the principle of induction, $P(n)$ is true for all $n \geq 0$. □

The theorem we originally set out to prove is restated below. With this invariant in hand, the proof is simple.

Theorem. *No sequence of moves transforms the board below on the left into the board below on the right.*

1	2	3	4
5	6	7	8
9	10	11	12
13	**15**	**14**	

1	2	3	4
5	6	7	8
9	10	11	12
13	**14**	**15**	

Proof. In the target configuration on the right, the total number of inversions is zero, which is even, and the blank square is in row 4, which is also even. Therefore, by Lemma 21, the target configuartion is unreachable. □

If you ever played with Rubik's cube, you know that there is no way to rotate a single corner, swap two corners, or flip a single edge. All these facts are provable with invariant arguments like the one above. In the wider world, invariant arguments are used in the analysis of complex protocols and systems. For example, in analyzing the software and physical dynamics a nuclear power plant, one might want to prove an invariant to the effect that the core temperature never rises high enough to cause a meltdown.

3.3 Unstacking

Here is another wildly fun 6.042 game that's surely about to sweep the nation! You begin with a stack of n boxes. Then you make a sequence of moves. In each move, you divide one stack of boxes into two nonempty stacks. The game ends when you have n stacks, each containing a single box. You earn points for each move; in particular, if you divide one stack of height $a + b$ into two stacks with heights a and b, then you score ab points for that move. Your overall score is the sum of the points that you earn for each move. What strategy should you use to maximize your total score?

As an example, suppose that we begin with a stack of $n = 10$ boxes. Then the game might proceed as follows:

stack heights										score
10										
5	5									25 points
5	3	2								6
4	3	2	1							4
2	3	2	1	2						4
2	2	2	1	2	1					2
1	2	2	1	2	1	1				1
1	1	2	1	2	1	1	1			1
1	1	1	1	2	1	1	1	1		1
1	1	1	1	1	1	1	1	1	1	1
								total score	=	45 points

Can you find a better strategy?

3.3.1 Strong Induction

We'll analyze the unstacking game using a variant of induction called *strong induction*. Strong induction and ordinary induction are used for exactly the same thing: proving that a predicate $P(n)$ is true for all $n \in \mathbb{N}$.

Principle of Strong Induction. Let $P(n)$ be a predicate. If

- $P(0)$ is true, and
- for all $n \in \mathbb{N}$, $P(0)$, $P(1)$, ..., $P(n)$ imply $P(n+1)$,

then $P(n)$ is true for all $n \in \mathbb{N}$.

The only change from the ordinary induction principle is that strong induction allows you to assume more stuff in the inductive step of your proof! In an ordinary induction argument, you assume that $P(n)$ is true and try to prove that $P(n+1)$ is also true. In a strong induction argument, you may assume that $P(0)$, $P(1)$, ..., $P(n-1)$, and $P(n)$ are *all* true when you go to prove $P(n+1)$. These extra assumptions can only make your job easier.

Despite the name, strong induction is actually no more powerful than ordinary induction. In other words, any theorem that can be proved with strong induction can also be proved with ordinary induction. However, an apeal to the strong induction principle can make some proofs a bit simpler. On the other hand, if $P(n)$ is easily sufficient to prove $P(n+1)$, then use ordinary induction for simplicity.

3.3.2 Analyzing the Game

Let's use strong induction to analyze the unstacking game. We'll prove that your score is determined entirely by the number of boxes; your strategy is irrelevant!

Theorem 22. *Every way of unstacking n blocks gives a score of $n(n-1)/2$ points.*

There are a couple technical points to notice in the proof:

- The template for a strong induction proof is exactly the same as for ordinary induction.
- As with ordinary induction, we have some freedom to adjust indices. In this case, we prove $P(1)$ in the base case and prove that $P(1), \ldots, P(n-1)$ imply $P(n)$ for all $n \geq 2$ in the inductive step.

Proof. The proof is by strong induction. Let $P(n)$ be the proposition that every way of unstacking n blocks gives a score of $n(n-1)/2$.

Base case: If $n = 1$, then there is only one block. No moves are possible, and so the total score for the game is $1(1-1)/2 = 0$. Therefore, $P(1)$ is true.

Inductive step: Now we must show that $P(1)$, ..., $P(n-1)$ imply $P(n)$ for all $n \geq 2$. So assume that $P(1)$, ..., $P(n-1)$ are all true and that we have a stack of n blocks. The

first move must split this stack into substacks with sizes k and $n-k$ for some k strictly between 0 and n. Now the total score for the game is the sum of points for this first move plus points obtained by unstacking the two resulting substacks:

$$\begin{aligned}
\text{total score} &= (\text{score for 1st move}) \\
&\quad + (\text{score for unstacking } k \text{ blocks}) \\
&\quad + (\text{score for unstacking } n-k \text{ blocks}) \\
&= k(n-k) + \frac{k(k-1)}{2} + \frac{(n-k)(n-k-1)}{2} \\
&= \frac{2nk - 2k^2 + k^2 - k + n^2 - nk - n - nk + k^2 + k}{2} \\
&= \frac{n(n-1)}{2}
\end{aligned}$$

The second equation uses the assumptions $P(k)$ and $P(n-k)$ and the rest is simplification. This shows that $P(1)$, $P(2)$, …, $P(n)$ imply $P(n+1)$.

Therefore, the claim is true by strong induction. □