Chapter 17

Proof by Contradiction

This chapter covers proof by contradiction. This is a powerful proof technique that can be extremely useful in the right circumstances. We'll need this method in Chapter 20, when we cover the topic of uncountability. However, contradiction proofs tend to be less convincing and harder to write than direct proofs or proofs by contrapositive. So this is a valuable technique which you should use sparingly.

17.1 The method

In proof by contradiction, we show that a claim P is true by showing that its negation $\neg P$ leads to a contradiction. If $\neg P$ leads to a contradiction, then $\neg P$ can't be true, and therefore P must be true. A contradiction can be any statement that is well-known to be false or a set of statements that are obviously inconsistent with one another, e.g. n is odd and n is even, or $x < 2$ and $x > 7$.

Proof by contradiction is typically used to prove claims that a certain type of object cannot exist. The negation of the claim then says that an object of this sort **does** exist. The existence of an object with specified properties is often a good starting point for a proof. For example:

Claim 54 *There is no largest even integer.*

> Proof: Suppose not. That is, suppose that there were a largest even integer. Let's call it k.
>
> Since k is even, it has the form $2n$, where n is an integer. Consider $k+2$. $k+2 = (2n)+2 = 2(n+1)$. So $k+2$ is even. But $k+2$ is larger than k. This contradicts our assumption that k was the largest even integer. So our original claim must have been true. □

The proof starts by informing the reader that you're about to use proof by contradiction. The phrase "suppose not" is one traditional way of doing this. Next, you should spell out exactly what the negation of the claim is. Then use mathematical reasoning (e.g. algebra) to work forwards until you deduce some type of contradiction.

17.2 $\sqrt{2}$ is irrational

One of the best known examples of proof by contradiction is the proof that $\sqrt{2}$ is irrational. This proof, and consequently knowledge of the existence of irrational numbers, apparently dates back to the Greek philosopher Hippasus in the 5th century BC.

We defined a rational number to be a real number that can be written as a fraction $\frac{a}{b}$, where a and b are integers and b is not zero. If a number can be written as such a fraction, it can be written as a fraction in lowest terms, i.e. where a and b have no common factors. If a and b have common factors, it's easy to remove them.

Also, we proved (above) that, for any integer k, if k is odd then k^2 is odd. So the contrapositive of this statement must also be true: (*) if k^2 is even then k is even.

Now, we can prove our claim:

> Suppose not. That is, suppose that $\sqrt{2}$ were rational.
>
> Then we can write $\sqrt{2}$ as a fraction $\frac{a}{b}$ where a and b are integers with no common factors.

Since $\sqrt{2} = \frac{a}{b}$, $2 = \frac{a^2}{b^2}$. So $2b^2 = a^2$.

By the definition of even, this means a^2 is even. But then a must be even, by (*) above. So $a = 2n$ for some integer n.

If $a = 2n$ and $2b^2 = a^2$, then $2b^2 = 4n^2$. So $b^2 = 2n^2$. This means that b^2 is even, so b must be even.

We now have a contradiction. a and b were chosen not to have any common factors. But they are both even, i.e. they are both divisible by 2.

Because assuming that $\sqrt{2}$ was rational led to a contradiction, it must be the case that $\sqrt{2}$ is irrational. □

17.3 There are infinitely many prime numbers

Contradiction also provides provides a nice proof of a classic theorem about prime numbers, dating back to Euclid, who lived around 300 B.C.

Euclid's Theorem: There are infinitely many prime numbers.

This is a lightly disguised type of non-existence claim. The theorem could be restated as "there is no largest prime" or "there is no finite list of all primes." So this is a good situation for applying proof by contradiction.

Proof: Suppose not. That is, suppose there were only finitely many prime numbers. Let's call them p_1, p_2, up through p_n.

Consider $Q = p_1 p_2 \cdots p_n + 1$.

If you divide Q by one of the primes on our list, you get a remainder of 1. So Q isn't divisible by any of the primes p_1, p_2, up through p_n. However, by the Fundamental Theorem of Arithmetic, Q must have a prime factor (which might be either itself or some smaller number). This contradicts our assumption that p_1, p_2,$\ldots p_n$ was a list of all the prime numbers. □

Notice one subtlety. We're not claiming that Q must be prime. Rather, we're making the much weaker claim that Q isn't divisible by any of the first n primes. It's possible that Q might be divisible by another prime larger than p_n.

17.4 Lossless compression

A final example concerns file compression. A file compression algorithm attempts to reduce the size of files, by transforming each input file to an output file with fewer bits. A **lossless** algorithm allows you to reconstruct the original file exactly from its compressed version, whereas a **lossy** algorithm only allows you to reconstruct an approximation to the original file. Or, said another way, a lossless algorithm must convert input files to output files in a one-to-one-manner, so that two distinct input files are never compressed to the same output file.

Claim 55 *A lossless compression algorithm that makes some files smaller must make some (other) files larger.*

> Proof: Suppose not. That is, suppose that we had a lossless compression algorithm A that makes some files smaller and does not make any files larger.
>
> Let x be the shortest file whose compressed size is smaller than its original size. (If there are two such files of the same length, pick either at random.) Suppose that the input size of x is m characters.
>
> Suppose that S is the set of distinct files with fewer than m characters. Because x shrinks, A compresses x to a file in S. Because no files smaller than x shrink, each file in S compresses to a file (perhaps the same, perhaps different) in S.
>
> Now we have a problem. A is supposed to be lossless, therefore one-to-one. But A maps a set containing $|S| + a$ files to a set containing $|S|$ files, so the Pigeonhole Principle states that two input files must be mapped to the same output file. This is a contradiction.

So, on the face of it, lossless file compression algorithms can't win. How do they work so well in practice? One secret is that compression algorithms can ensure that file sizes never increase much. If a file would increase in size, the algorithm stores the original version unchanged, preceded with a one-bit marker. This bounds the potential damage if we encounter a "bad" input file.

The second secret is that commonly-occurring files are not created at random but have definite patterns. Text files contain natural language text. Digitized images contain values that tend to change gradually. Compression algorithms are tuned so that common types of files shrink. The fact that some files might get bigger isn't a serious practical problem if those files are unlikely to occur on your disk.

17.5 Philosophy

Proof by contradiction strikes many people as mysterious, because the argument starts with an assumption known to be false. The whole proof consists of building up a fantasy world and then knocking it down. Although the method is accepted as valid by the vast majority of theoreticians, these proofs are less satisfying than direct proofs which construct the world as we believe it to be. The best mathematical style avoids using proof by contradiction except when it will definitely result in a much simpler argument.

There is, in fact, a minority but long-standing thread within theoretical mathematics, called "constructive mathematics," which does not accept this proof method. They have shown that most of standard mathematics can be re-built without it. For example, the irrationality of $\sqrt{2}$ can be proved constructively, by showing that there is an error separating $\sqrt{2}$ from any chosen fraction $\frac{a}{b}$.