Math 79SI Notes

Mason Rogers

3 Proof By Contrapositive

3.1 Background

In the last section, we introduced the concept of an implication as a statement of form 'if P then Q', where P is a list of hypotheses and Q is a list of conclusions. We then discussed proving implications directly by assuming that the hypotheses are true and deducing that the conclusions must be true as well. However, we also saw a few instances in which this approach did not work easily. Sometimes, it may seem tempting to manipulate the (failure of the) *conclusions* and work towards the (failure of the) *hypotheses*. Over the next two weeks, we will introduce two similar strategies that allow us to do this in a logically sound way: proof by contrapositive and proof by contradiction. This week, we will study proof by contrapositive, for which we assume that the conclusions of a statement are false and show that its hypotheses must be false as well. Next week, we will study proof by contradiction, for which we assume that the hypotheses of a statement hold but its conclusions do not and show that an absurdity arises.

Given an implication 'if P then Q', we define the *contrapositive* statement to be 'if *not* Q then *not* P'. Any implication is logically equivalent to its contrapositive,[1] so proving either proves both. Thus, if proving the original statement directly is difficult, we may try proving its contrapositive instead.

We can think of this in two ways: (I) we take the contrapositive of the original statement and prove it directly, or (II) we assume that the conclusion of the original statement is *false*, then show that at least one hypothesis must be *false* as well. For instance, consider Proposition 2.1, which states that if a and b are odd numbers, their product ab is odd. We proved this directly in section 2.2, but we could also prove it by contrapositive by assuming that a product of two integers a and b is even and proving that at least one of a and b is even.

As a warning, while a statement (if P then Q) is equivalent to its contrapositive (if not Q then not P), it is *not* equivalent to its *converse* (if Q then P) or its *inverse* (if not P then not Q). For example, consider the statement for a real-coefficient polynomial p that if p has odd degree, then p has a real root, a special case of which was proven in Corollary 2.11. The converse of this statement is that if p has a real root then p must have odd degree, which is clearly false.

[1] Why? Suppose $P \implies Q$ is true and Q is false. If P were true then $P \implies Q$ would fail, so P must be false. Thus, if $P \implies Q$ holds, so does its contrapositive. Repeating the same argument while switching the roles of the contrapositive and the original completes the proof.

It is common to make and prove *biconditional* statements that can be viewed as containing both 'forwards' and 'backwards' implications. These are useful to prove the equivalence of different sets of conditions, each of which may be easier to verify or proceed from in different circumstances. Such statements are worded as 'P if and only if Q' and mean 'if P then Q *and* if Q then P', i.e. both $P \implies Q$ and $Q \implies P$. Sometimes, they are abbreviated with the bidirectional implication arrow ' $\iff$ ' or the shorthand 'iff', as in '$P \iff Q$' or 'P iff Q'. For example, the statements 'a triangle is equilateral if and only if all of its angles are $60°$' and 'a prime number p is a factor of ab if and only if p is a factor of a or p is a factor of b' are both true biconditional statements.

When confronted with biconditional statements, you must prove an implication and its converse or inverse.[2] In light of our warning not to mistake the converse or inverse for the contrapositive when proving a statement by contrapositive, we make a similar warning here. It is important not to prove a statement and its contrapositive when trying to prove a biconditional statement, as this amounts to proving the same implication twice.

We conclude with a few related example statements and consider whether or not they hold biconditionally. If $\triangle ABC$ and $\triangle A'B'C'$ are congruent (with A corresponding to A', B to B', and C to C'), the following properties hold:

- $AB = A'B', BC = B'C', CA = C'A'$ (SSS condition)
- $AB = A'B', BC = B'C', m(\angle B) = m(\angle B')$ (SAS condition)
- $m(\angle A) = m(\angle A'), AC = A'C', m(\angle C) = m(\angle C')$ (ASA condition)
- $AB = A'B', BC = B'C', m(\angle C) = m(\angle C')$ (SSA condition)

The contrapositives of these statements therefore also clearly hold; if any of the conditions fails, the triangles are not congruent. It is also helpful to know which of those four conditions are sufficient to deduce that $\triangle ABC$ is congruent to $\triangle A'B'C'$, i.e. which statements have true converses. As you may remember from high school geometry, the SSS, SAS, and ASA conditions imply congruence of the triangles, but the SSA condition does not. This is because of an issue involving acute and obtuse angles, depicted in Figure 1.

These examples are helpful to keep in mind to remember that while statements are equivalent to their contrapositives, they are not always biconditional.

3.2 Examples

Let us begin with some simple examples from earlier in the course to see when proof by contrapositive might be convenient. Recall that in our discussion following Problem 1.2, we mentioned that if a natural number n factors as ab, then at least one of the factors must be at least $\sqrt{n}$. We will prove a slight generalization of this statement by contrapositive.

Proposition 3.1. If $x, y, z \in \mathbb{R}_{>0}$ satisfy $z \leq xy$, then $x \geq \sqrt{z}$ or $y \geq \sqrt{z}$.[3]

[2] Note that the converse and inverse of a statement are contrapositives of each other and therefore can be substituted.

[3] The symbol $\mathbb{R}_{>0}$ is a common shorthand for the set of positive real numbers.

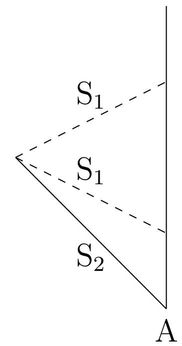

Figure 1: Observe that by reflecting S$_1$ about the horizontal axis, we can construct two triangles which satisfy the SSA property but are not congruent.

Proof. Beginning with $z \leq xy$ does not provide a clear path to bounding x or y, so we attempt a proof by contrapositive. The contrapositive of the original statement is 'if $x, y, z \in \mathbb{R}_{>0}$ satisfy $x < \sqrt{z}$ and $y < \sqrt{z}$, then $z < xy$'. Note that the inequalities involved are now *strict* and all quantities are nonnegative, so we have

$$xy < \sqrt{z} \cdot \sqrt{z} = z,$$

as desired. □

Another example we encountered in a previous week is Proposition 2.2. We revisit it armed with our new technique of proof by contrapositive.

Proposition 3.2. If the product ab of two natural numbers a and b is odd, then a and b are both odd.

Proof. As we already realized in the previous section, it is difficult to prove this statement directly. However, the contrapositive statement 'if a and b are natural numbers and at least one of them is even, then ab is even' is easily proven. Without loss of generality, let a be even, so $a = 2k$ for some natural number k. Then $ab = 2kb$, and because kb is an integer, ab must be even. □

Note that this proof along with the proof of Proposition 2.1 is also a proof of the biconditional statement 'the product ab of natural numbers a and b is odd *if and only if* a and b are both odd'. We offer one more example of a statement that is difficult to prove directly but easy to prove by contrapositive.

Proposition 3.3. For $n \in \mathbb{N}$, if n^2 is odd then n is odd, and if n^2 is even then n is even.

Proof. We will first prove that if n^2 is odd then n is odd. If we were to prove this directly, our assumption would give us that $n^2 = 2k + 1$ for some integer k. However, it is not straightforward to continue from here, as we cannot easily discern much about $n = \sqrt{2k+1}$. Instead, we aim to prove the contrapositive, which states that if n is even then n^2 is even. To do so, simply let $n = 2k$ for some integer k and observe that $n^2 = 4k^2$, which is clearly even.

The other statement is similar; beginning with $n^2 = 2k$ results in the same problem as before, as $\sqrt{2k}$ is not readily simplified. We therefore turn to the contrapositive approach, leaving us to prove that if n is odd then n^2 is odd. Analogously to the first case, let $n = 2k+1$ for some integer k, and observe that $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$, which is clearly odd. $\square$

We see that sometimes, even when a direct proof may seem impossible, a proof by contrapositive falls out almost immediately. Of course, it usually takes more work to achieve the proof by contrapositive, as in the next example.

Recall that in the lead-up to Proposition 1.2, in which we factored the fifth Fermat number, we used a result without proof that if a number of form $2^m + 1$ is prime, then $m = 2^n$ for some nonnegative integer n. We now prove this.

Proposition 3.4. If m is a natural number such that $2^m + 1 > 2$ is prime, then $m = 2^n$ for some nonnegative integer n.

Proof. The direct approach would be to let p be a prime number of form $2^m + 1$ for $m \in \mathbb{N}$, then show that $m = 2^n$ for some nonnegative integer n. This seems rather difficult, though, as we do not have an immediate strategy for taking a prime and generating information about its form. To prove the contrapositive, we assume that m is *not* of form 2^n and aim to deduce that $2^m + 1$ is *not* prime. We *do* have a tool for going from information about a number's form and showing that the number is composite: factoring.

Suppose that m is *not* of form 2^n for some nonnegative integer n. Since $1 = 2^0$, m must be a natural number greater than 1. By the existence of a prime factorization for m, we can therefore write $m = 2^\ell(2k+1)$, where ℓ is a nonnegative integer and $k \in \mathbb{N}$. Our goal is now to show that $2^{2^\ell(2k+1)} + 1$ factors nontrivially. This will generalize the familiar factorization $x^3 + 1 = (x+1)(x^2 - x + 1)$.

Observe that

$$2^m + 1 = 2^{2^\ell(2k+1)} + 1 = \left(2^{2^\ell}\right)^{2k+1} + 1.$$

It may not be immediately apparent how to factor this expression, but letting $x = 2^{2^\ell}$, we have $\left(2^{2^\ell}\right)^{(2k+1)} + 1 = x^{2k+1} + 1$. For any $x \in \mathbb{R}$, this expression factors as

$$x^{2k+1} + 1 = (x+1)\left(\sum_{i=0}^{2k} (-x)^i\right),$$

which can be proven by mathematical induction on k, a topic we will study in the coming weeks. In the case at hand, we substitute 2^{2^ℓ} for x to obtain

$$2^m + 1 = \left(2^{2^\ell} + 1\right)\left(\sum_{i=0}^{2k} \left(-2^{2^\ell}\right)^i\right).$$

Note that since the left side and the first factor on the right side are positive, the second factor on the right side must also be positive, so all 3 terms are in $\mathbb{N}$. All that remains to check is that this factorization is nontrivial, i.e. that neither factor is 1. It suffices to check that one factor is neither 1 nor 2^m+1 itself.[4] This is easy to verify, as $1 < 2^{2^\ell}+1 < 2^m+1$ because $m = 2^\ell(2k+1)$ for some $k \geq 1$ by our hypothesis on m. Thus, we have found a nontrivial factorization of 2^m+1, concluding the proof of the proposition. $\square$

While taking the contrapositive of the previous statement made the proof manageable, it certainly did not allow the proof to fall out immediately. Over time, you will develop the skill of discerning when a direct proof or a proof by contrapositive might be most appropriate. It is a good rule of thumb to estimate whether the assumptions or conclusions of a statement seem easier to manipulate, but if the approach you settle on fails, you can always try the other.

We now consider a few more challenging examples which we have not encountered earlier in the course. The first comes from calculus and characterizes local extrema of functions.

Proposition 3.5. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function. If f achieves a local maximum at $x = a$, then $f'(a) = 0$.

Proof. The intuition for this statement seems to follow the line of a direct proof. Given a plot of the graph of f, we identify local extrema by finding points where the tangent line to the graph of f is horizontal. Thus, it makes sense that f' would vanish at those points. However, the condition of achieving a local maximum is difficult to manipulate concretely, so we instead prove the contrapositive.

For this problem, we will use the limit definition of the derivative without worrying too much about formalities involving limits. Suppose that $f'(a) \neq 0$; we aim to prove that f does not achieve a local maximum at $x = a$. From the limit definition of the derivative, we have

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{h\to 0} \frac{f(a)-f(a-h)}{h} = c \neq 0.$$

First, suppose $c > 0$. Then for some $h > 0$ and all h' satisfying $0 < h' < h$, we have that

$$\frac{f(a+h')-f(a)}{h'} > 0,$$

and consequently that $f(a+h') > f(a)$. In particular, for arbitrarily small $h > 0$, we can choose $h' < h$ such that $f(a+h') > f(a)$. Thus, f does not achieve a local maximum at $x = a$.

If $c < 0$, we use a similar argument with the other expression we have for c. For some $h > 0$ and all h' satisfying $0 < h' < h$, we have that

$$\frac{f(a)-f(a-h')}{h'} < 0,$$

and consequently that $f(a) < f(a-h')$. Again, for arbitrarily small $h > 0$, we can choose $h' < h$ such that $f(a-h') < f(a)$, so f does not achieve a local maximum at $x = a$. $\square$

[4] Why? Suppose $n = ab$ for some $a, b, n \in \mathbb{N}$. The statement 'if $a \neq n$ then $b \neq 1$' is the contrapositive of (and thus equivalent to) 'if $b = 1$ then $a = n$', which is true by the assumption that $n = ab$. Thus, if a is neither 1 nor n, then neither a nor b is 1.

A similar argument proves the case that the extremum in question is a local minimum. In either case, the condition of having a local extremum at a point is more difficult to manipulate than the condition of having nonzero derivative, hence why the contrapositive approach is useful.

3.3 Exercises

1. Write the contrapositives of the following (true) statements:

   (a) Choose $q \in \mathbb{R}$. If q is a rational number, then q has an eventually repeating decimal expansion.

   (b) Choose $x \in \mathbb{R}$. If $x \geq 0$, then $x = y^2$ for some real number y.

   (c) For all p in the set of prime numbers, there exists a larger prime q.

2. The following statements were proven for functions $f : X \to Y$ and $g : Y \to Z$ as a part of the previous section's exercises. Write their converses and prove or disprove those.

   (a) If f and g are injective functions, then $g \circ f$ is an injective function.

   (b) If f and g are surjective functions, then $g \circ f$ is a surjective function.

   Moreover, consider the related statement 'if f is injective and g is surjective, then $g \circ f$ is bijective'. Prove or disprove it, then write its converse and prove or disprove that.

3. Prove that if $x, y, z \in \mathbb{R}$ satisfy $z = x + y$, then $x \geq \frac{z}{2}$ or $y \geq \frac{z}{2}$.

4. Choose $x, y \in \mathbb{R}$, and suppose that $xy \notin \mathbb{Q}$. Prove that $x \notin \mathbb{Q}$ or $y \notin \mathbb{Q}$.

5. Write the converse to the statement proven in the previous exercise, and prove it or disprove it.

6. Mathematicians call primes of form $2^n - 1$ *Mersenne Primes* after Marin Mersenne, a French friar who studied them in the early 17th century. Mersenne primes are important for their connection to perfect numbers (a concept in number theory) and random number generation, among other reasons. Prove that if $2^n - 1$ is prime, then n must be prime. (*Hint: you may use without proof that* $x^{ab} - 1 = (x^a - 1)(x^b + x^{2b} + \ldots + x^{(a-1)b})$ *for any* $x \in \mathbb{R}$ *and any* $a, b \in \mathbb{N}$. *We will be able to prove this later with mathematical induction, but you may be able to see already why this may be true.*)

7. Prove Proposition 3.8.

8. Look up and summarize a proof of the Steiner-Lehmus theorem. What is its broad approach? Does it use trigonometry or algebra, or does it simply use elementary geometry? Does it use the contrapositive or establish a contradiction?

3.4 Further Reading

CPZ Chapter 3.3	Introduction to proof by contrapositive with simple examples.
CPZ Chapter 4	More advanced direct proofs and proofs by contrapositive mixed together.
CPZ Chapter 2.6	More about biconditionals and converses if you are confused.
Sherri R. Gardner "A Variety of Proofs of the Steiner-Lehmus Theorem"	The author discusses and classifies many classical proofs of the Steiner-Lehmus theorem. dc.etsu.edu/cgi/viewcontent.cgi?article=2332&context=etd
Wikipedia entry for Mersenne Primes	What Mersenne Primes are and why mathematicians care. https://en.wikipedia.org/wiki/Mersenne_prime
Wikipedia entry for Primality Tests	There are many different primality tests with various efficiencies and sensitivities. https://en.wikipedia.org/wiki/Primality_test

3.5 Activity

Until the 19th century, many calculus textbooks featured 'proofs' of statements that are incorrect. Notable examples include the statements that (I) any function can be represented locally by a power series and (II) continuous functions of real numbers must be differentiable outside a set of isolated points.

In the 19th century, Karl Weierstrass provided a famous family of examples of functions that are continuous everywhere but differentiable nowhere, clearly disproving statement (II). One example of Weierstrass' functions is $f(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} \cos(13^n \pi x)$.[5] We will omit the definitions of differentiability and continuity as well as a proof that $f(x)$ has the desired properties for now. But if you plot the first few terms of the series for f, you will see that the graph of f features small but rapid oscillations, confirming our *intuition* for what an everywhere continuous, nowhere differentiable function might look like.

A counterexample to statement (I) is given by

$$f(x) = \begin{cases} 0 & x \leq 0 \\ e^{-\frac{1}{x^2}} & x > 0, \end{cases}$$

which is a useful function in many areas of math because of how smoothly and quickly it decays near $x = 0$. We again omit the proof that this example contradicts statement (I), but plotting $f(x)$ indicates that it behaves very smoothly near 0, even though the behavior on either side of 0 is very different. We will study these functions more in future exercises.

The reason that the false proofs had stood for so long had to do with an unclear notion of a function, which was originally defined as an 'analytic expression', whatever that means. Today, we will try to construct a better definition of 'function' that includes the above examples and appropriately summarizes what our sense of a function is.

[5] The infinite sum must 'converge' at each x to make sense, a notion which you may recall from calculus experience. We will introduce a *definition* of convergence in the coming weeks. For now, we will accept that this sum converges. More information is available in the further reading section 2.4.

Step 1 Everybody independently construct a list of as many objects that they think should be counted as functions as possible.

Step 2 Everybody independently construct what they believe to be a mathematically precise definition of a function that can handle all of their test cases.

Step 3 Volunteers share their definitions and test cases. If one student's definition does not handle another student's test case, we all can discuss whether the definition needs to be expanded or the test case rejected. Afterwards, we will reach a definition of function that the class agrees on.

Step 4 Test the class definition of function against the following test cases:

- $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$
- $f : \mathbb{R} \to \mathbb{R}, f(x) = 27\sin(3x^{12}) + 463x^3 - 1$
- $f : \mathbb{R} \to \mathbb{R}, f(x) = 1$ if $x \in \mathbb{Q}$, $f(x) = 0$ otherwise
- $f : \mathbb{N} \to \mathbb{N}, f(n) =$ [the index of the first appearance of n in π]
- $f : \{\text{dog, cow, elephant}\} \to \{\text{dog food, grass, plants}\}, f([\text{animal}]) = [\text{food it eats}]$

Additionally, compare the class definition to the definition of a function $f : X \to Y$ as a subset $A_f \subset X \times Y$ such that every element of X is the first component of exactly one ordered pair in A_f.

3.6 Further Examples

We give another example of a proof by contrapositive from number theory. This result can be stated and proven rather quickly using the language of *modular arithmetic*, which we are not assuming but will explore in future exercises. Instead, we state it in terms of division and remainders.

Proposition 3.6. If a natural number n leaves remainder 3 after dividing by 4, then n is not a sum of two perfect squares.

Proof. We begin with a preliminary claim about what remainders perfect squares can leave after dividing by 4.

Claim. For $m \in \mathbb{N}$, m^2 leaves remainder 0 or 1 after dividing by 4.

Proof of Claim. We split into the cases m is even and m is odd and proceed directly. Suppose m is even, so $m = 2k$ for some $k \in \mathbb{Z}$. Then $m^2 = 4k^2$, which leaves remainder 0 after dividing by 4. Now suppose m is odd, so $m = 2k+1$ for some $k \in \mathbb{Z}$. Then $m^2 = 4(k^2+k)+1$, which leaves remainder 1 after dividing by 4. This covers all cases for m, so we are done. □

Now we return to proving the contrapositive of the statement in the proposition. Suppose that n can be written as $k^2 + \ell^2$ for some $k, \ell \in \mathbb{N}$. By the claim, k^2 and ℓ^2 each leave remainder 0 or 1 after dividing by 4, giving four cases: $k^2 = 4a, \ell^2 = 4b$, $k^2 = 4a+1, \ell^2 = 4b$, $k^2 = 4a, \ell^2 = 4b+1$, and $k^2 = 4a+1, \ell^2 = 4b+1$. In each respective case, the remainder of $n = k^2 + \ell^2$ after dividing by 4 is 0, 1, 1, and 2. This exhausts all possible cases, so we conclude that if n leaves remainder 3 after dividing by 4, n cannot be written as the sum of two squares. $\square$

This result is perhaps more exciting in light of several other results due to Euler. First, if p is a prime which leaves remainder 1 after dividing by 4, then p necessarily can be written as a sum of two perfect squares. Second, if a natural number n and factor a can be written as a sum of two perfect squares, so can the quotient $\frac{n}{a}$. These results, combined with the technique of mathematical induction, the fundamental theorem of arithmetic, and the fact that $2 = 1^2 + 1^2$, yield the following theorem.

Theorem 3.7. Let $n > 1$ be a natural number, and let $\prod_{i=1}^{k} p_i^{e_i}$ be the unique prime factorization of n. Then n can be written as a sum of two perfect squares if and only if none of the p_i leaves remainder 3 when dividing by 4.[6]

We offer one more famous historical example from geometry of a proof that is difficult by contrapositive, but perhaps not even possible directly. As a warm-up, consider the following statement, which you may have proven in a high school geometry class.

Proposition 3.8. Let $\triangle ABC$ be an isosceles triangle with $\overline{AB} = \overline{AC}$. Then the angle bisectors of $\angle ABC$ and $\angle ACB$ are congruent.

Informally, this follows from the symmetry of the triangle; you are welcome to work out the details if you are up for a challenge. The converse to this statement, a useful tool for proving that triangles are isosceles, is much more challenging to prove:

Theorem 3.9 (*Steiner-Lehmus*)**.** Let $\triangle ABC$ be a triangle such that the bisectors of $\angle ABC$ and $\angle ACB$ are congruent. Then $\triangle ABC$ is isosceles with $\overline{AB} = \overline{AC}$.

A number of proofs (as well as an anthology and classification of them) are available in the further reading section 3.4. There is still debate over whether a direct proof exists, as most available proofs utilize either the contrapositive or the contradiction approach.

We now move on from discussing proof by contrapositive specifically and focus our attention on proving biconditional statements. We have already discussed the hypothetical utility of being able to generate large primes quickly, and it is similarly useful to be able to check quickly if large numbers are prime. One of the most crude (in that it is slow) but effective (in that it definitively determines whether a number is prime) primality tests is known as trial division.

Proposition 3.10. Choose $n \in \mathbb{N}$ such that $n > 1$. Then n is prime if and only if for all $m \in \mathbb{N}$ such that $1 < m \leq \sqrt{n}$, m does not divide n.

Proof. Let us unpack this statement before proving it. The phrase 'if and only if' signals to us that we must prove two implications: that n being prime implies n has no divisors less than or equal to $\sqrt{n}$ other than 1, and that n having no divisors less than or equal to $\sqrt{n}$ other than 1 implies n is prime.

[6] If you are interested in seeing proofs for the results leading up to this theorem, this theorem itself, or theorems like this in general, you may want to take a course in elementary number theory or algebra.

The former is easy, as n being prime implies that n has no divisors other than 1 and itself by the definition of a prime number. Since $n > 1$, we have $\sqrt{n} < n$, so n does not have any divisors m satisfying $1 < m \leq \sqrt{n}$.

The latter implication is more challenging, and we will prove its contrapositive. Suppose that n is not prime; we aim to show that n admits a factor m such that $1 < m \leq \sqrt{n}$. Since $n > 1$, n is necessarily composite. We thus write $n = ab$ for $a, b \in \mathbb{N}$ such that neither $a = 1$ nor $b = 1$. As proven in Proposition 3.1, either $a \geq \sqrt{n}$ or $b \geq \sqrt{n}$. Without loss of generality, suppose $a \geq \sqrt{n}$. Then

$$b = \frac{n}{a} \leq \frac{n}{\sqrt{n}} = \sqrt{n},$$

so $b \leq \sqrt{n}$. Thus, b is a factor of n satisfying $1 < b \leq \sqrt{n}$, concluding the proof of the second implication.

With both implications now proven, we are finished the proof of the proposition. $\square$

This method can be improved upon the recognition that we must only check for *prime* factors p of n such that $1 < p \leq \sqrt{n}$, as any factor of n larger than 1 must have a prime factor.

When proving a biconditional statement, it is helpful to consider each implication separately, as we did above. This helps avoid the aforementioned pitfalls of mistaking the converse, inverse, and contrapositive, as well as making the argument more readable.