What is a Proof? [Next]

Suppose you are a detective working the case of the murder of Unfortunate Ursula. Your investigation thus far has turned up the following:

The suspects in the death of Ursula are Suspicious Sam, Illegal Irene, and Violent Victor. Either Sam or Irene has a key to Ursula's house, but you don't know which. The gardener said that either Victor or Sam had been lurking in the garden the morning before the murder, but this turned out to be false. However, whoever had the key had been lurking in the garden. If Irene killed Ursula, then she entered through a window and wouldn't have had a key to Ursula's house. If Victor killed Ursula, he did so with Irene's help.

Sitting down with these clues, you reason as follows:

<table>
<thead>
<tr>
<th>(a) Either Sam or Irene has a key to Ursula's house.</th>
<th>[From the second sentence]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b) It is not true that either Victor or Sam had been lurking</td>
<td>[From the third sentence]</td>
</tr>
</tbody>
</table>
in the garden.

(c) So, Victor had not been lurking in the garden, and Sam had not been lurking in the garden. [From (b)]

(d) Whoever had the key had been lurking in the garden. [From the fourth sentence]

(e) If Sam had the key, Sam had been lurking in the garden. [From (d)]

(f) Sam didn’t have the key. [From (c) and (e)]

(g) Irene had the key. [From (a) and (f)]

(h) If Irene killed Ursula, then she didn’t have the key. [From the fifth sentence]

(i) Irene didn’t kill Ursula. [From (g) and (h)]

(k) If Victor killed Ursula, then Irene also killed Ursula. [From the sixth sentence]

(l) Victor didn’t kill Ursula. [From (j) and (k)]

(m) Either Sam killed Ursula, or Victor killed Ursula, or Irene killed Ursula. [From the first sentence]

(n) Sam killed Ursula. [From (j) and (l) and (m)]

By stringing together a series of deductive inferences, you are able to move from the information you had to begin with to the identity of the killer.

The above is an example of a proof. In a proof, one moves in steps from one piece of information to another, by making deductively valid steps. Since deductive reasoning is transitive, these steps can then be put together to form an inferential path from some starting information to some final conclusion.

Proofs thus provide another method of determining validity of arguments. If there is a proof starting from some premises and ending with a certain conclusion, then those premises imply that conclusion.

But why do we need another method of determining validity? After all, the method of truth tables already gives us a wholly general and fully mechanical method for making all such determinations. Here are five reasons for supplementing the method of truth tables with a proof system:

- In the long run, probably the central reason is that a proof system offers a syntactic method for checking argument validity, while the method of truth tables offers a semantic method. A syntactic method for checking arguments appeals only to syntactic features of the argument -- namely, the shapes of the symbols in the sentences and the patterns in which those symbols are arranged. A semantic method, on the other hand, appeals to the meanings of the symbols. The advantage of a syntactic proof system is that it can be used by someone with no knowledge of what any of the symbols mean. So, for example, a Martian who knew absolutely nothing about English could still learn to do proofs in English (given a syntactic proof system for English -- we're not going to give such a proof system for English (just for our formal logic), though, because the complexities of English are too great) just by learning to manipulate symbols in the right way. More immediately usefully, a computer can perform proofs in a syntactic proof system without needing to know the meanings of any of the symbols.

As we set up our proof system, we'll see how it is a syntactic system and can be used without knowing what the symbols mean. The contrast with semantic methods of testing arguments is unfortunately not as clear yet as it could be, since the method of truth tables (although technically a semantic method, because it appeals to the semantic notion of truth-in-an-interpretation) is given a largely syntactic implementation (it's not difficult, for example, to write a computer program to generate truth tables and use them to check arguments). When we come to more complex logics later, the method truth tables will no longer be adequate to their semantic complexities, and we'll see the gap between proof-based syntactic proof methods of checking implication and truth-in--an-interpretation semantic methods of checking implication much more clearly.

- Truth tables become prohibitively long when many sentence letters are involved. The proof given above involves at least seven sentence letters, which means a truth table of 128 lines. Even a moderately complex proof can easily involve 20 different sentence letters, which means a truth table of 1,048,576 lines. A proof with 250 sentence letters would require a truth table with more lines than there are subatomic particles in the universe. However, a proof involving 250 sentence letters can often take no more than around 1000 steps. (We won't be doing any 1000 step proofs, by the way.) A proof system thus offers a considerable increase in efficiency.

- There is no equivalent to the method of truth tables for more complex logics that we will soon be considering. We'll return to this point once we introduce those logics.

- The method of truth tables can be used to check an argument only when you already know what the conclusion of the argument will be (this is a slight oversimplification). The method of proofs, on the other hand, allows you to take some
premises and start reasoning from them, and see where you end up (as we did in the example above). In real life, one rarely knows the answer before one gets started, so the exploratory character of the method of proofs is better suited to conditions of uncertainty.

- This last one is a bit harder to pin down, but there is a sense in which a proof gives an explanation for the validity of an argument in a way that a truth table does not. By seeing an inferential path from the premises to the conclusion laid out, one sees how the premises necessitate the truth of the conclusion, and how and in what way each premise is involved, in a way that one does not when one simply lays out a truth table.

A Formal Proof System

We now want to give a precisely defined proof system for our formal logic. We will start with the thought that a proof consists of a list of lines, where each line contains both (a) some claim, in the form of some sentence, and (b) a justification for making that claim. The idea is that as we proceed through a proof we must always be able to explain why we are in a position to assert whatever we are asserting.

We thus need a list of permissible justifications for lines of proofs. You can think of designing a proof system as being like designing a game. Consider, for example, the game of chess. A particular game of chess can be thought of as a series of board positions, where a board position is a specification of the location of each piece on the board. Thus the initial position of a chess game might be:

- White rook on a1 and white knight on b1 and ... and black knight on g8 and black rook on h8.

and a typical ending position might be:

- White king on e1 and black rook on a1 and black queen on a2 and black king on g8.

If you have these two positions, and all the intervening positions, then you have a full description of that game of chess. However, not just any series of board positions counts as a game of chess. We cannot, for example, just put together the two board descriptions above to make a game, because no rule in chess allows one to remove most of the opposing player's pieces and place their king in check.

To characterize the general notion of a game of chess, then, we need some rules of transition, which tell us how we can move from one board state to another legally. These rules are the movement rules for chess, which stipulate which piece can move where, depending on the setup of the board.

An Aside: The rules of transition for chess have a modular structure since they are given via movement rules for each piece, any of which can be appealed to at any time. However, it needn't have been that way. We could have had global transition rules, like "if there is a white piece in each of the four corners of the board, move all bishops to the center of the board". In fact, not all actual chess rules are modular -- the rule for castling involves the positions (and histories) of two pieces. Keeping the chess rules largely modular, however, simplifies the game, since it can then be thought of as involving movements of individual pieces, rather than entire transformations of the board.

Not just any rule can be appealed to at any point, however. Given, for example, the initial position of a chess game, we cannot use the rook movement rule to justify a transition to a position in which a rook has advanced from a1 to a6, because there is an intervening pawn. In addition to appealing to a rule, we need to show that use of that rule is appropriate in the circumstances.

The conditions of appropriateness of rule application can be as elaborate as we want. In chess, most rules have only the appropriateness condition that a piece can only move through open squares (although the knight is partially exempted from even that condition). However, some rules are more elaborate. A king cannot move into check, which requires checking ways that the movement rules could be applied to the result of applying the king movement rule to the current board position. Castling can occur only if neither the king nor the rook have moved, which requires looking not just at the current board position, but at all the previous board positions as well.

Putting these ideas together, we will want our logical proofs to be a sequence of lines, each of which gives:
- a sentence (a "position" in the proof game)
- a proof rule (a "movement rule" in the proof game)
- one or more earlier lines in the proof (a justification of the appropriateness of the proof rule)

What we need to do next, then, is say what the proof rules are -- what the rules are of the game of constructing proofs.
We might ask first why we need to give an explicit list of proof rules. Consider some of the transitions made in the proof we started with, such as:

- Either Sam or Irene had a key to Ursula's house.
- Sam didn't have the key.
- So, Irene had the key.

or:

- If Sam had the key, Sam had been lurking in the garden.
- Sam had not been lurking in the garden.
- So, Sam didn't have the key.

It's tempting to say here that it's just obvious that these are valid transitions, and that there's no need to encode a series of rules telling us what transitions are valid. We are not going to give in to this temptation, however. Instead, we are going to give a rather small list of rules which give the permissible moves. Setting up our proof system in this way will likely be a source of great frustration, because it will mean that there are moves which you know are valid (in the sense that you know they really follow from what you have so far), but which you can't make immediately (because there is no proof rule for that move) and you can't figure out how to make in a sequence of moves (just because you can't see it -- there will always be a suitable sequence of moves; it's just a matter of finding it). So it's worth taking some time to justify the decision to set things up this way.

- First, it's not always obvious which transitions are permissible and which are not. Suppose I reason like this:
  ♦ Someone loves someone
  ♦ Everyone loves someone who loves someone
  ♦ So, everyone loves everyone.
  or like this:
  ♦ The sequence x₁, x₂, x₃, ... converges at 0.
  ♦ Therefore, the sum of this sequence is finite.
  or like this:
  ♦ If I do X, then an omniscient being knows that I do X.
  ♦ If an omniscient being knows that I do X, then I must do X.
  ♦ So, if I do X, then I must do X.
  We are unlikely to get wide consensus on whether these are valid forms of reasoning. Rather than requiring ourselves to spend lots of time arguing about each form of reasoning we decide to use, it's more convenient to settle ahead of time on a collection of inferential transitions we all agree on.

- Second (and this point is largely a variant on the first), every inferential principle we agree to is a risk we assume in our proof system. Each new proof rule we endorse is one more place where error can creep into the proof system, since if the proof rule we endorse is not in fact a valid form of reasoning, it will lead to incorrect results. Caution thus gives us a reason to adopt as few rules as we can, so that the possibility of error is minimized.

- Third, there is a certain intellectual and aesthetic virtue in starting with as sparse an apparatus as possible, and seeing how all the pieces are eventually built up from that minimal beginning. We are going, eventually, to show how to erect the entire architecture of mathematics on this humble beginning, and it's worth seeing how little we need to get things started.

- Fourth, part of the point of working with a proof system is learning to reason with an artificially constricted set of tools. It's thus similar to the intellectual exercise of translating English into the restricted semantic tools of a logical language, or of finding a way to implement a task in a low-level computer language. This isn't the only point of the proof system, but it is one of them.

- Fifth, we will eventually want to move from producing proofs in the proof system to producing proofs about the proof system (establishing results, for example, about what it can and cannot prove). If we keep our proof system parsimonious, it will be much easier to produce proofs about it, because it will be smaller and more compact. If, on the other hand, we throw in every inference procedure we can think of, we'll have a big, bulky proof system, which will be very inconvenient to produce proofs about, since we'll have to consider many different cases corresponding to the many different inference rules. Prudent reasoning thus dictates that we go through a bit of inconvenience now in order to make things easier on our future selves.
Sixth, it turns out that that inconvenience won't actually be that great. We'll quickly start to add new proof rules to the system, without actually adding to the complexity of the system, by showing that certain new rules can be added by definition since they can always be reproduced using a combination of the original rules (this will be something like defining a macro). More on this later.

So, there we have six reasons for restricting ourselves to a small group of proof rules. With that many reasons, the position must surely be correct. (Bertrand Russell once said that you should always be wary of people who offer many reasons for a conclusion, since if the reasons were any good, they'd only need one.)

The Proof Rules, At Last

Finally we are ready to start introducing the proof rules. Like the movement rules for chess, the proof rules for our system will be highly modular. While the movement rules for chess typically control the behaviour of a single piece, the proof rules for our system will typically relate to a single sentential connective. The general idea is that for each sentential connective there will be two rules. Each of these rules for a given connective will be invoked in connection with a sentence in the proof whose main connective is that connective. One of the two rules will be an elimination rule, used to get rid of the connective where it has already appeared in the proof, and the other will be an introduction rule, used to bring the connective in to the proof where it had not been before. (We won't be able to hold exactly to this plan, and deviations from it will be noted as we go along.)

The First Rule: Conjunction Elimination

We'll begin with the two conjunction rules, which are the most straightforward in the system. The first conjunction rule is the rule of conjunction elimination, which we'll abbreviate &E, for reasons which are hopefully obvious.

Conjunction elimination allows one to make the inferential move from $P \& Q$ to $P$, or from $P \& Q$ to $Q$. It thus captures the thought that if a conjunction is true, both of its conjuncts must also be true (this feature of a conjunction is, naturally, also captured in the truth table for the conjunction -- the truth of the conjunction guarantees that one is on the first line of the truth table (since that is the only T), and the first line of the truth table forces the truth of both $P$ and $Q$).

Let's now look at the way conjunction elimination works when it is used in a proof. Consider the following couple of lines (these two lines don't make up a whole proof, for reasons we'll get to soon, but just a proof fragment).

| (1) $P \& Q$ | &E, 1 |

The left-hand column of the proof (fragment) contains, on each line, some sentence (the sentence being established on that line). The right-hand column contains two items: an appeal to a proof rule to justify the sentence appearing to the left, and a citation of some earlier line (or lines) of the proof showing that the proof rule is appropriate to use in this context.

On the second line we write the sentence $P$, and justify writing this sentence by appealing (in the right-hand column) to the rule of conjunction elimination (&E). Conjunction elimination allows us to move from a conjunction to one of its conjuncts, so in order to use &E here, we need the proof already to contain some conjunction having $P$ as one of its conjuncts. Thus the right-hand column also includes a citation of line 1 of the proof. Line 1 contains the sentence $P \& Q$, which is indeed a conjunction having $P$ one of its conjuncts, so the use of &E is appropriate here.

One way of thinking of proof rules is as input-output filters. Given an input of the form of a conjunction -- that is, of the form $\phi \& \theta$ -- the rule yields an output of the form of either of the two conjuncts -- that is, of the form $\phi$ or of the form $\theta$.

We can use the following format (which will be our standard format for introducing new proof rules) to illustrate the input-output structure of the &E rule:

$$
\&E: \quad \phi \& \theta \quad \text{or} \quad \phi \& \theta
$$

Here the sentence above the line gives the input conditions for the proof rule, and the sentence below the line gives the output result of applying that proof rule to the input. Because &E offers a choice of outputs, there are two input-output formats for the rule -- you can choose to invoke either one when you use the rule.

Before moving on, let's cover some questions about what's been said so far:

**Question #1**: Why are there Greek letters in the input-output specification of the &E rule?
**Answer:** The Greek letters are serving to represent arbitrary sentences. In the first example, we applied &E to the simple conjunction \( P \land Q \), but &E can be applied to any conjunction whatsoever. So, for example, the following is a valid use of &E:

\[
\begin{align*}
(1) & (P \rightarrow Q) \land \neg(P \leftrightarrow (R \lor \neg S)) \\
(2) & \neg(P \leftrightarrow (R \lor \neg S)) & \text{&E, 1}
\end{align*}
\]

In this example, we have:
\( \Phi = (P \rightarrow Q) \)
\( \Theta = \neg(P \leftrightarrow (R \lor \neg S)) \)
so we are, using &E, justified in concluding \( \Theta \) — that is, in concluding \( \neg (P \leftrightarrow (R \lor \neg S)) \). If you want a fancy name for the Greek letters, they're called **metavariables**. (A name which will probably make more sense once we talk about both variables and metalanguages later.)

In general, as long as the main connective of the sentence is a conjunction, the two conjuncts can be as complicated as we want, and we can still apply &E.

**Question #2:** What if the main connective is not a conjunction, but there’s still a conjunction somewhere in the sentence? Can I still use &E?

**Answer:** No. Definitely not. One of the most important features of all of the proof rules is that they can only be applied when their connective is the main connective of the input sentence. Without this feature, the proof system will lead immediately to false results.

Let’s look at an example. Suppose I do the following:

\[
\begin{align*}
(1) & \neg (P \land Q) \\
(2) & \neg P & \text{&E, 1}
\end{align*}
\]

(I’ve put this proof fragment in red to show that it is incorrect.) The application of &E in line 2 is incorrect, because the main connective of line 1 (the input line to the &E rule) is a negation, not a conjunction. Since the main connective is a negation, conjunction elimination cannot be used on this line.

A little thought will show that we need this restriction on &E to avoid getting bad results. Think about the (fallacious) proof above, in which we “derived” \( \neg P \) from \( \neg (P \land Q) \). If we consider an interpretation in which:
- \( P = T \)
- \( Q = F \)
we will have \( \neg (P \land Q) = T \), and \( \neg P = F \), so the “inference” would be taking us from a true premise to a false conclusion. Since a valid argument can never take us from a true premise to a false conclusion, this must not be a valid argument. Since it’s not a valid argument, we don’t want our proof system to permit it.

If you’re still uncertain about when &E can be used and when it can’t, try the following quiz. Figure out which uses of &E are legal, and then check your choices against the answers.

| (A) | \( \neg P \land Q \) | \( \neg P \) &E, 1 |
| (B) | \( P \rightarrow (Q \land R) \) | \( P \rightarrow R \) &E, 1 |
| (C) | \( P \land R \rightarrow R \) | \( P \rightarrow R \) &E, 1 |
| (D) | \( P \leftrightarrow R \land (R \leftrightarrow P) \) | \( R \leftrightarrow P \) &E, 1 |
| (E) |
Question #3: If I can't use &E on $\neg (P \land Q)$ because $\land$ is not the main connective, how do I make inferences from $\neg (P \land Q)$?

Answer: Using rules for negation, since negation is the main connective in $\neg (P \land Q)$. We'll get to those rules soon.

Question #4: Does the conjunction need to come immediately before the use of &E?

Answer: No. As long as there is a conjunction somewhere earlier in the proof, you can use &E to derive either of its conjuncts. Thus the following is a perfectly legal use of &E:

1. $P \land Q$
2. $R \lor S$
3. $Q \&E, 1$
4. $R \&E, 2$

Make sure when you do this, however, that the line you cite in conjunction with &E (as providing the required input for the &E rule) is the line containing the conjunction, and not just the immediately prior line.

Question #5: If there is more than one line earlier in the proof with a conjunction, can I use &E on any of the earlier conjunctions I want?

Answer: Yes. Suppose you've got the following:

1. $P \land Q$
2. $R \land S$
3. $Q \&E, 1$
4. $R \&E, 2$

You can then continue the proof by applying &E to line 1 or by applying &E to line 2:

1. $P \land Q$
2. $R \land S$
3. $Q \&E, 1$
4. $R \&E, 2$

Question #6: Can I derive both conjuncts from a single conjunction?

Answer: Yes, as long as you derive each on its own line. The following is a valid proof fragment:

1. $P \land Q$
You cannot, however, derive both conjuncts from a conjunction on the same line. What would it mean to do this, since you can only put one sentence on each line?

**Question #7**: Can I derive the same conjunct more than once from the same conjunction?

**Answer**: Yes, as long (of course) as you do it on different lines. As we’ll see later, there are times when it is in fact useful to do this.

**Question #8**: All of these proof fragments you’ve given so far have had conjunctions on their first line without any justification given in the right-hand column. What’s going on with that? Where did the conjunction come from?

**Answer**: Excellent question. Those lines have come out of nowhere, which is not acceptable in a proof, since every line needs some justification. Obviously, though, we’re not going to get these lines using &E. (Well, we could get them from &E if we had even earlier in the proof an even bigger conjunction of which our current conjunction was one conjunct. But then we’d have to worry about where that conjunction came from. We could appeal to &E again, but clearly an infinite regress is starting here.) So we’re going to need to add another rule to our proof system, which is....

### The Zeroeth Rule: Assumption

Go back for a minute to our analogy between a chess game and a proof system. Suppose that I’ve given you all the movement rules for all the pieces in a game of chess. Do you now know how to play the game of chess? No – there’s still one crucial piece of information you’re lacking (actually, there are two crucial pieces of information you’re lacking – we’ll come to the second in a bit). You don’t know how the board is to be set up at the beginning of the game. The actual movement rules of chess are compatible with any number of starting positions (exercise for the reader: figure out how many), and these different starting positions will lead to very different games of chess.

So to give a complete description of the game of chess, you also need to specify the starting position of the board. Similarly, when specifying a proof system, you need to specify the starting position of the proof. If our rules are thought of as input-output filters, then they are not going to be able to produce any results unless they are first given some input(s) to filter to some outputs. So we need a way of giving initial inputs – of getting the proof started.

There is, however, an important difference between the starting position of chess and the starting position of a proof. All chess games start with the same position – the major pieces along each of the back ranks, and the pawns in the next rank forward. Proofs, however, do not all start from the same position. The starting position of a proof can be thought of as the information we already know, from which we want to reach conclusions. So, in the little “murder mystery” scenario we considered at the beginning of this discussion, the “starting position” was the information you-as-detective had already gathered, from which you wanted to infer the identity of the killer. However, we obviously don’t want to start every single proof with a list of information about Sam or Irene having a key, and Victor or Sam lurking in the garden, and so on. We still want to be able to reason things through if the facts of the case are different, or if we’re engaged in some completely different bit of reasoning (some mathematics, or a political argument, or some philosophy).

What we need, then, is a special rule that allows us to introduce our “starting position” in whatever way we want. This rule will be the **A**, or Assumption, rule. The use of the **A** rule is very simple. This rule can be used to introduce any sentence whatsoever. At the beginning of a proof, you simply write on the left side whatever premise you want, and then cite the **A** rule to justify introducing that premise. You can also introduce as many premises as you want, although they must all be introduced before you start any other work on the proof (this last requirement isn’t really necessary, but it adds greatly to the readability of a proof).

So to return to our first application of the &E rule above, we would now have:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) P &amp; Q</td>
<td>A</td>
</tr>
<tr>
<td>(2) P</td>
<td>&amp;E, 1</td>
</tr>
</tbody>
</table>

The first rule now has a justification – the **A** rule, which tells us that the claim P & Q is not derived from something we already knew, but is just one of our starting assumptions. We could also have a proof with more than one starting assumption:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) P &amp; Q</td>
<td>A</td>
</tr>
<tr>
<td>(2) (¬Q ∨ R) &amp; T</td>
<td>A</td>
</tr>
<tr>
<td>(3) Q</td>
<td>&amp;E, 1</td>
</tr>
</tbody>
</table>
Here the first two lines are both assumptions, and then we begin applying our derivation rules (so far, only &E) to the starting information given in the assumptions.

A few questions about the A rule, and answers to them:

Question #1: Isn't the A rule absurdly powerful? Suppose I'm trying to prove some conclusion P – why would I bother using any of the other rules if I can just write down P on a line and cite the A rule? I'll be able to finish any proof in a single step.

Answer: It's quite true that you can reach any conclusion you want instantly using the A rule. However, this isn't as useful as it might seem at first. You have to keep in mind here what the goal of a proof is. When you're doing a proof, you're not just trying to construct any old proof that ends with the desired conclusion. Instead, you want to show that the conclusion follows from reasonable assumptions.

Suppose, for example, that I say that I can produce a proof that you owe me $1,000,000. I then proceed to produce the following:

| (1) You owe me $1,000,000. | A |

and ask you to hand over the cash. You're quite unlikely to be swayed by what I've done. My proof is perfectly valid, of course – but it's only a convincing argument if you've got reason to believe the assumption. And, of course, if you had reason to believe the assumption (namely, that you owe me $1,000,000), there would be no need to construct a proof deriving the conclusion (namely (again), that you owe me $1,000,000). Suppose instead that I construct this proof:

| (1) You owe me $1,000,000 and the moon orbits the earth. | A |
| (2) You owe me $1,000,000. &E, 1 |

Again, this is a perfectly valid proof, but one which is unlikely to sway you, since once again you'd have to already believe the conclusion in order to find the premise acceptable.

For a proof to be interesting, then (as opposed to just valid), it has to move from assumptions you have some reason to accept to some conclusion. This means that you can't start with just any old assumptions and produce an interesting proof (although you can produce a valid one). In the context of a logic class, the way this will typically play out is that you will be given some set of assumptions which you are allowed to use in deriving the desired result. Outside the logic class in the real world the situation is somewhat less straightforward, but in general there will be some stuff that you already believe, and you'll be interested in seeing how that stuff can be extended to new beliefs using a proof. It's of no interest in this process to look at proofs whose assumptions you don't believe, because the validity of such proofs will do nothing to encourage you to believe their conclusions.

Question #2: What line numbers do I cite as inputs when I use the A rule?

Answer: None. The A rule doesn't require inputs, since it's used to set up the starting position of the proof, which starting position will then be used as inputs for subsequent rules as the proof continues.

Question #3: Actually, I can't think of any more questions – this rule seems pretty straightforward.

Answer: Right then – on to the next.

The Second Rule: Conjunction Introduction
Conjunction elimination gives us a way to move from a line with a conjunction to a line without that conjunction.

Brief Digression: Notice, by the way, that &E won't necessarily take us from a line with a conjunction to a line without any conjunction. Consider the following proof:

| (1) P & (Q & R) | A |
| (2) Q & R | &E, 1 |

Here the second line – obtained through the use of &E – does contain a conjunction. However, it does not contain the very same conjunction which made possible the application of &E to the first line.
Conjunction elimination, as I was saying, gives us a way to move from a line with a conjunction to a line without that conjunction. Having a way to get rid of conjunctions, what we need now is a way to introduce new conjunctions where we want them. That way will, probably not surprisingly, be the rule of conjunction introduction, or &I.

A slight new wrinkle is necessary for introducing &I, though. Suppose we were to try to set up &I on analogy with &E. We might try something like this:

\(\text{P} \quad A \)
\(\text{P} \& \text{Q} \quad \&I, 1 \)

Here the rule takes us – as we wanted – from a sentence without a conjunction to a sentence with a conjunction. However, a little thought will show that this is not the rule we want. The problem is that the inference this rule just permitted is not a good one. \(P\) does not in fact imply \(P \& Q\), as we can easily see by considering the following interpretation:

\(P = T\)
\(Q = F\)

which makes the premise (line 1) true and the conclusion (line 2) false.

The rough problem here is that \(P \& Q\) contains more information than just \(P\), so we can’t "safely" proceed from \(P\) to \(P \& Q\), because that additional portion of information might be false (as it – the sentence \(Q\) – is in the interpretation we just gave). Logical implication flows down the information hill – a valid proof can take one to a state of less information, or to a different state of the same information, but it can never take you to a state of more information. That’s the price one pays for deductive certainty.

The downhill flow of implication presents no problems for &E, because &E moves one from a high information state (knowing that both \(\Phi\) and \(\Theta\) are true) to a low information state (knowing only that \(\Phi\) is true, or knowing only that \(\Theta\) is true). But in &I, our desired ending point – \(\Phi \& \Theta\) – is a high information state, and is thus harder to get to. We’re going to need to start, as it were, further uphill to be able to get there.

But where to start? We already saw that starting at \(P\) won’t get us to \(P \& Q\), and it’s pretty easy to see that starting at \(Q\) is no better:

\(\text{Q} \quad A \)
\(\text{P} \& \text{Q} \quad \&I, 1 \)

We could try starting at \(P \& Q\):

\(\text{P} \& \text{Q} \quad A \)
\(\text{P} \& \text{Q} \quad \&I, 1 \)

but that’s pretty silly – the &I rule won’t actually accomplish anything for us that way. I suppose we could try starting with the conjunction in the other order:

\(\text{Q} \& \text{P} \quad A \)
\(\text{P} \& \text{Q} \quad \&I, 2 \)

but this doesn’t look much like a conjunction introduction rule either, and it’s a bit mysterious where we’d get \(Q \& P\), if we didn’t have it as a premise. You might try starting with some sentence built out of other connectives – ¬, v, →, ↔ – I’ll leave it as an exercise for the reader to show that this doesn’t work either.

So what do we do? In all likelihood you’re banging your head on the wall with frustration at the delay in getting to the obvious answer, so we’ll cut to the chase now. What we’ll do is require two inputs to the &I rule, instead of just one. The two inputs will be the two halves of the conjunction we want to introduce. Deriving \(P \& Q\) will, then, work like this:

\(\text{P} \quad A \)
\(\text{Q} \quad A \)
\(\text{P} \& \text{Q} \quad \&I, 1, 2 \)
Notice that two lines are cited as inputs to the &I rule, instead of just one (as with the &E rule). Both P and Q are below P & Q on the information slope, but by having both of them as starting points, we have enough informational power to derive P & Q validly.

Here, then, is the official formulation of the &I rule (given, as with the &E rule, in input-output format)

\[
\begin{array}{ll}
\phi & \\
& & \\
& & \\
\phi & \& & \Theta \\
& & \\
& & \\
\phi & \& & \Theta \\
\end{array}
\]

The fact that there are two sentences above the line shows that two inputs to the &I rule are needed, which in conjunction lead to the conclusion below the line.

Let's look at a few questions about the &I rule, and then go on to look at some more elaborate examples.

**Question #1**: Why are there Greek letters in the input-output specification of the &I rule?

**Answer**: They're metavariables. Please see the answer to the first question about the &E rule.

**Question #2**: Can I only combine sentence letters into a conjunction with the &I rule?

**Answer**: No, you can combine any two sentences you want into a conjunction using the &I rule. Thus all three of the following inferences are valid:

(1) P A
(2) Q A
(3) P & Q &I, 1,2

(1) Q ≡ T A
(2) P → (R & S) A
(3) (Q ≡ T) & (P → (R & S)) &I, 1,2

(1) P & Q A
(2) R & S A
(3) (P & Q) & (R & S) &I, 1,2

**Question #3**: Why do those extra parentheses show up when &I is applied, in the examples you just gave? For example, why do we have Q ∨ R in line 2 of the first example, and then parentheses around Q ∨ R in line 3?

**Answer**: Our notational conventions allow us to drop outermost parentheses, so we can write (Q ∨ R) (the proper form) as Q ∨ R. However, once Q ∨ R is made into half of a conjunction, it is no longer outermost, so we need parentheses around it (to avoid ambiguity about the scopes of connectives).

**Question #4**: Can I use &I to combine two lines into a conjunction even if neither of those lines has a conjunction in it?

**Answer**: Sure. Any two sentences can be combined using &I. The first example in the answer to the previous question shows two lines, neither of which contains a conjunction, being combined into a conjunction with &I.

**Question #5**: Can I use &I to combine two lines if they already have conjunctions in them?

**Answer**: Yes. Any two sentences can be combined using &I. The third example in the answer to Question #2 shows two lines, both of which are conjunctions, being combined into an even larger conjunction with &I.

**Question #6**: Do I have to put the conjuncts into the conjunction in the same order in which they appeared in the proof?

**Answer**: No. Both of the following proofs are proper:

(1) P A
(2) Q A
(3) P & Q &I, 1,2
The conjuncts can be assembled in either order – it's entirely up to the constructor of the proof.

Question #7: I noticed that in the previous example you cited the lines in the reverse order (2,1 instead of 1,2) when you were putting the conjuncts together in a different order. Is this required?

Answer: No. I've written it that way to make it clearer what's going on, and will usually do it that way in proofs I write, but the strict rules of the system place no constraint on the order in which the rules are cited.

Question #8: Do the two lines providing the two conjuncts have to be next to each other, or next to the line resulting from &I?

Answer: No. As with &E, the input lines can appear anywhere earlier in the proof.

Question #9: Can I use the &I rule on a line multiple times?

Answer: Yes. The following is a proper proof:

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>Q</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>P &amp; Q</td>
</tr>
<tr>
<td>5</td>
<td>Q &amp; R</td>
</tr>
<tr>
<td>6</td>
<td>(P &amp; Q) &amp; (R &amp; S)</td>
</tr>
</tbody>
</table>

In this proof, Q on line 2 serves once as input to &I on line 4, and then again as input to &I on line 5.

Question #10: Can a single line provide both inputs to &I?

Answer: Yes. The following is a proper proof:

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>P &amp; P</td>
</tr>
</tbody>
</table>

Believe it or not, this is actually useful from time to time.

Question #11: If I've got three lines that I want to form into a big conjunction, can I use &I on all three of them at once, like this:

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>Q</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>P &amp; Q &amp; R</td>
</tr>
</tbody>
</table>

Answer: Strictly speaking, no (as you might have guessed, from the fact that the "proof" was in red).

Question #12: So, how do I form a conjunction of all three?

Answer: Like this:

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>Q</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>P &amp; Q</td>
</tr>
<tr>
<td>5</td>
<td>(P &amp; Q) &amp; R</td>
</tr>
</tbody>
</table>

or like this:
Notice that the conclusions of these two proofs are not the same, although they are equivalent.

**Question #13**: But that's really a nuisance, isn't it?

**Answer**: Yes, and it's worse than you probably think, because it's also a pain going the other direction. If I want to infer R from P & (Q & R), I have to use &E twice:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise</th>
<th>Rule</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P &amp; (Q &amp; R)</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Q &amp; R</td>
<td>&amp;E, 1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>&amp;E, 2</td>
<td>2</td>
</tr>
</tbody>
</table>

It really starts to add up, if you're doing a lot of building up and breaking apart of big conjunctions.

To some extent, such inconveniences are just part of having a tightly constrained proof system. Some things will look obviously true, but it will take a surprising amount of effort to prove them. On the whole this is good, because it helps us catch those cases where something looks obviously true but in fact is false. However, while it's good on the whole, it's still bad here. Later we'll start saving time by adopting the notational convention of dropping parentheses in extended conjunctions, and thus writing P & Q & R rather than either (P & Q) & R or P & (Q & R). Once we do this, we'll then allow &I to be used to derive the three-part conjunction in a single step from three inputs, and also allow &E to be used to extract any of the three conjuncts.

**Some Sample Proofs [Next]**

Now that we've got two rules in place (well, three if you count the A rule), we can do some real proofs. None of the results we will get will be terribly substantive, of course, since we can't manipulate anything other than conjunctions yet. But we can at least see how the two rules can work together.

We'll start by proving an instance of the commutivity of conjunctions. To say that a conjunction is commutative is to say that it doesn't matter what order the conjuncts come in. Of course, you have different sentences depending on the order of the conjuncts, but those different sentences are equivalent, and thus in some sense say the same thing.

Suppose, then, we have the sentence P & Q. We want to show that we can swap the order of the conjuncts -- that P & Q is equivalent to Q & P. We can show that P & Q implies Q & P with the following proof:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise</th>
<th>Rule</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P &amp; Q</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P</td>
<td>&amp;E, 1</td>
<td>3,2</td>
</tr>
<tr>
<td>3</td>
<td>Q</td>
<td>&amp;E, 1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Q &amp; P</td>
<td>&amp;I, 3,2</td>
<td></td>
</tr>
</tbody>
</table>

Let's go through this proof a step at a time just to make sure it's clear what's going on. In the first line, we state our assumption. We're assuming that P & Q is true, and trying eventually to show from that assumption that Q & P is true. In the second and third lines, we use the rule of &E to extract each of the two conjuncts from the conjunction on the first line. The &E rule can be applied as many times as we want to a given conjunction, so we are free to extract both conjuncts. We then use the &I rule on the two conjuncts we have just extracted, and reassemble them into a conjunction -- but in the opposite order. &I allows us to put the two conjuncts in either order, so we are free to conclude Q & P. (We could also have concluded P & Q, but this would obviously have been a less interesting conclusion, given our starting point.)

**Question**: Did I have to first extract P using &E, and then extract Q using &E, or could I have do it in the other order?

**Answer**: The other order also would have worked. There is often more than one way to do the same proof.

The broad structure of the proof, then, is to use an elimination rule to break our starting assumption into small pieces, and then use an introduction rule to take those small pieces and put them together into the shape we desire. This kind of structure will be extremely common in our proofs, so you should start getting in the habit of looking for it.
We've now shown that \( P \land Q \) implies \( Q \land P \), but we haven't yet shown that the two are equivalent, because we haven't shown that \( Q \land P \) implies \( P \land Q \). Doing so requires another proof:

\[
\begin{align*}
(1) & \quad Q \land P & \text{A} \\
(2) & \quad Q & \text{\&I, 1} \\
(3) & \quad P & \text{\&I, 1} \\
(4) & \quad P \land Q & \text{\&I, 2,3}
\end{align*}
\]

Obviously, this proof is extremely similar to the proof that \( P \land Q \) implies \( Q \land P \). These two proofs taken together suffice to show that \( P \land Q \) and \( Q \land P \) are equivalent.

Have we now shown that conjunction is commutative? No. It's important to be clear on what is and is not shown by a proof. We've shown that \( P \land Q \) is equivalent to \( Q \land P \), but we haven't (for example) shown that \( (R \lor \neg S) \land \neg T \) is equivalent to \( \neg T \land (R \lor \neg S) \), which is another instance of the commutivity of conjunction. Of course, we could easily whip off a pair of proofs establishing this equivalence as well. They'd look very much like the proofs we just did -- one of them would be:

\[
\begin{align*}
(1) & \quad (R \lor \neg S) \land \neg T & \text{A} \\
(2) & \quad R \lor \neg S & \text{\&E, 1} \\
(3) & \quad \neg T & \text{\&E, 1} \\
(4) & \quad \neg T \land (R \lor \neg S) & \text{\&I, 3,2}
\end{align*}
\]

and the other would go the other direction. This proof would establish the equivalence of \( (R \lor \neg S) \land \neg T \) with \( \neg T \land (R \lor \neg S) \), but there would remain an infinite number of other pairs of conjunctions to show equivalent in order to establish the general claim that conjunction is commutative.

A particular proof can only establish implication facts between (or among) particular sentences. To prove a general claim about implication facts (like the claim that conjunction is commutative), we'd have to produce a proof about proofs (for example, a proof that a proof of the right sort could be produced for any appropriate pair of conjunctive sentences). Much later we'll learn some techniques for producing proofs about proofs (metaproofs, we might call them), but for now we'll limit ourselves to the informal observation that it's obvious that we can always make a proof along the above lines.

So, we've now established the amazing result that \( P \land Q \) is equivalent to \( Q \land P \). You're probably at this point wondering why it's necessary to have an elaborate proof system in order to produce proofs of stunningly obvious results. It's a fair question; hopefully seeing the applications of the proof system as we develop it more will help answer it. Let's try, then, proving another result. This time we will prove an instance of the claim that conjunction is associative -- that it doesn't matter what order you group conjuncts in a conjunction (it's this result which justifies us in adopting the notational convention of dropping internal parentheses in a conjunction). We'll prove the particular instance that \( (P \land Q) \land R \) is equivalent to \( P \land (Q \land R) \). This will, as before, require two proofs, one in each direction. First we have:

\[
\begin{align*}
(1) & \quad (P \land Q) \land R & \text{A} \\
(2) & \quad P \land Q & \text{\&E, 1} \\
(3) & \quad R & \text{\&E, 1} \\
(4) & \quad P & \text{\&E, 2} \\
(5) & \quad Q & \text{\&E, 2} \\
(6) & \quad Q \land R & \text{\&I, 5,3} \\
(7) & \quad P \land (Q \land R) & \text{\&I, 4,6}
\end{align*}
\]

Notice that this proof, like the proof that \( P \land Q \) implies \( Q \land P \), follows the pattern of breaking the assumption into component pieces with an elimination rule, and then reassembling those pieces using an introduction rule.

**Question #1:** Could I have used \&E on \( P \land Q \) to separate \( P \) and \( Q \) before I used \&I on \( (P \land Q) \land R \) to separate \( R \)?

**Answer:** Yes. The order of these doesn't matter; line 3 of the current proof could be moved after lines 4 and 5.

**Question #2:** Could I have performed the two \&I steps at the end of the proof in the other order, linking up \( P \) before linking \( Q \) and \( R \)?

**Answer:** No -- this would have brought you back right where you started. You would have gotten the following:
Question #3: How do you know where the parentheses go when you form line 7?

Answer: The parentheses must go around Q & R, rather than around P & Q, because Q & R is one of the inputs (from line 6). If this causes confusion, try explicitly placing outermost parentheses around all lines, and things should go more clearly.

To complete the equivalence result, we need a (very similar) proof running in the other direction:

\[(1) \quad P \land (Q \land R) \\
(2) \quad P \\
(3) \quad Q \land R \\
(4) \quad Q \\
(5) \quad R \\
(6) \quad P \land Q \\
(7) \quad (P \land Q) \land R \land I, 6,5 \]

Let's do one more proof with the conjunction rules before moving on. This time we will combine associativity and commutivity by rearranging and regrouping a conjunction. We will show that \((P \land Q) \land (R \land S)\) implies \(S \land (R \land (Q \land P))\).

\[(1) \quad (P \land Q) \land (R \land S) \land A \\
(2) \quad P \land Q \land &E, 1 \\
(3) \quad P \land &E, 2 \\
(4) \quad Q \land &E, 2 \\
(5) \quad R \land S \land &E, 1 \\
(6) \quad R \land &E, 5 \\
(7) \quad S \land &E, 5 \\
(8) \quad Q \land P \land &I, 4,3 \\
(9) \quad R \land (Q \land P) \land &I, 6,8 \\
(10) \quad S \land (R \land (Q \land P)) \land &I, 7,9 \]

Notice that after we finish the disassembling of the original conjunction (on line 7), we have to put together the new conjunction from the inside out. We start with the conjunction of smallest scope, and work our way out to the conjunction of largest scope. When applying the elimination rules, on the other hand, we always start with the connective of largest scope and work our way inward to the connective of smallest scope. (In fact, the definition of the rules forces us to do this, since the elimination rule can only be applied to the connective of largest scope in the input.)

One More Element of Proofs

I mentioned earlier that knowing the movement rules for chess, even when combined with knowing the starting position for chess, still wasn't enough to fully characterize the game of chess. To see that there's still something more to know, consider the game of suicide chess. In suicide chess, the goal of the game is to force the other player to checkmate you. The movement rules for all the pieces are the same, and the starting position of the board is the same. If a game of suicide chess were written down (in the form of a sequence of board positions), you wouldn't be able to tell that it was, in fact, suicide chess and not regular chess.

The crucial piece of information you'd be missing, which would distinguish the game of suicide chess from the game of regular chess, is the identity of the winner. If the final board position has the black king checkmated by white, and the winner of the game is white, then the game was suicide chess. If, on the other hand, the final board position has the black king checkmated by white, and the winner of the game is black, then the game was suicide chess. To the movement rules and starting position, then, we need to add a specification of victory conditions, in order to fully characterize the game of chess.

The movement rules of chess correspond to the proof rules of our proof system, and the starting position to the A rule. What, then, corresponds to the victory conditions? So far, the answer is “nothing”, but we will now change that. Intuitively, the victory condition of a proof is successfully proving what you want to prove. But notice that nothing in the proof specifies what you want to prove. We are now going to change that by adding to the proof system the idea of a Show line. A Show line is a line in a proof which does not assert some piece of information as known, but rather declares an intention to prove that something is true.
Show lines are used in proofs by introducing (in the left-hand column) a line containing the word Show followed by some sentence (the sentence you want to prove). No rule is cited in the right-hand column for a Show line -- this is because Show lines do not assert anything as true, and thus require no justification (you can want to prove anything you want, without justifying yourself -- you only need justification when you claim to have succeeded in proving something).

Consider, for example, the proof we gave earlier that P & Q implies Q & P. Now that we've added Show lines, our proof would begin like this:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P &amp; Q</td>
</tr>
<tr>
<td>2</td>
<td>Show Q &amp; P</td>
</tr>
<tr>
<td>3</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

After setting out the starting assumptions of the proof (here just one assumption), we introduce a Show line indicating the goal of the proof. Notice that in addition to not needing a justification, the content of the Show line cannot serve as input to any proof rule. Thus the following would not be a legal proof:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>Show Q &amp; P</td>
</tr>
<tr>
<td>3</td>
<td>Q</td>
</tr>
<tr>
<td>4</td>
<td>Q &amp; P</td>
</tr>
</tbody>
</table>

We can't use the content of line 2 here because we haven't actually shown it to be true yet -- we've just declared that we want to show it true. If we then tacitly assume that it is true by using it as an input to the rule &E, and then use the output of that rule to derive what we said (in the Show line) that we wanted to prove, we're simply begging the question -- arguing in a circle. It's perhaps hard to see how anyone could make this mistake, but it can be a surprisingly seductive one, especially when you're stuck on a proof and can't see how to proceed, and especially when you're working with some of the more complex proof forms we'll be getting to later.

A proper proof, then, proceeds by using a Show line to declare an intention to derive some result, and then proceeds to use proof rules without using the content of the Show line as input. Once we have derived (that is, obtained on some line (a non-Show line)) the sentence we set out to Show, we cancel the Show line by crossing it out, thereby indicating that we have done what we set out to do. Thus the complete proof that P & Q implies Q & P would look this this:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P &amp; Q</td>
</tr>
<tr>
<td>2</td>
<td>Show Q &amp; P</td>
</tr>
<tr>
<td>3</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>Q</td>
</tr>
<tr>
<td>5</td>
<td>Q &amp; P</td>
</tr>
</tbody>
</table>

We will think of Show lines as the output mechanism for the entire proof -- in order to extract information as established by a proof, it must appear on a cancelled Show line.

**Question #1:** You said earlier that one of the virtues of a proof system is that you didn't have to know what the answer was before you started deriving. But if you have to have a Show line indicating what you want to prove, doesn't that eliminate that advantage? Consider the "murder mystery" case again -- how would we know what to put on the Show line until we have worked through the proof and discovered who the murderer is?

**Answer:** Well, yes. It does eliminate that advantage. But there's a way around it. If we're in an exploratory situation, where we don't know what follows from some assumptions and just want to find out, we can just put a blank Show line after the assumptions, derive for a while, and then when we reach some interesting conclusion go back and write that conclusion into the Show line (and then, of course, cancel the Show line, since we will have derived that result).

Really at this point the whole device of Show lines is somewhat of a technicality. As far as what we've done so far is concerned, we could just as well have stayed with the idea that every line in a proof counts as a result established by that proof. However, things are going to get more complicated as we go, and soon we're going to need a distinction between conclusory lines and internal-workings lines.

**Question #2:** So if I derive some sentence \( \phi \) in my proof, but I don't have an earlier line saying Show \( \phi \), then I don't count as having proved \( \phi \)?
**Answer:** Technically speaking, that's right. However, you can fix it up easily by going back and adding the appropriate Show line.

**Question #3:** Does the Show line have to come before the line in the proof deriving the content of the Show line? Would this be a valid proof?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) P &amp; Q</td>
<td>A</td>
</tr>
<tr>
<td>(2) P</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(3) Show P</td>
<td></td>
</tr>
</tbody>
</table>

Answer: Well, it's a pretty arbitrary decision, but we'll go ahead and insist that Show lines must appear before the deriving of what they tell you to show, if only to improve readability of the proofs.

**Question #4:** Can I have more than one Show line in a proof?

**Answer:** This is a good question. The answer is "yes", although for the moment it's hard to see why we would want to. In theory, you can make any declarations as you want about your intentions to derive results. Thus you could have a proof like the following:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) P &amp; Q</td>
<td>A</td>
</tr>
<tr>
<td>(2) Show Q &amp; P</td>
<td></td>
</tr>
<tr>
<td>(3) Show P</td>
<td></td>
</tr>
<tr>
<td>(4) P</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(5) Show Q</td>
<td></td>
</tr>
<tr>
<td>(6) Q</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(7) Q &amp; P</td>
<td>&amp;I, 5,3</td>
</tr>
</tbody>
</table>

This proof would produce Q & P, P, and Q all as outputs. However, for reasons we'll get to soon, some care is required when we have multiple show lines.

When a proof has multiple Show lines, it's really best thought of as multiple interconnected proofs. Each Show line can be thought of as invoking a subproof. We will use indentation to mark subproof structure (the book uses bracketing lines off to the side, but those are a pain to produce). With subproof structure indicated, the proof we just gave would be:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) P &amp; Q</td>
<td>A</td>
</tr>
<tr>
<td>(2) Show Q &amp; P</td>
<td></td>
</tr>
<tr>
<td>(3) Show P</td>
<td></td>
</tr>
<tr>
<td>(4) P</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(5) Show Q</td>
<td></td>
</tr>
<tr>
<td>(6) Q</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(7) Q &amp; P</td>
<td>&amp;I, 5,3</td>
</tr>
</tbody>
</table>

The Show on line 3 triggers a subproof which is then carried out on line 4. The Show on line 5 also triggers a subproof, which is carried out on line 6. (Subproofs can, and generally will, take more than a single line to complete, but we won't see this until we add more complexity to the proof system.)

The crucial point here (and we'll come back in more detail to this later) is that subproofs must be self-contained. Results that are derived within the subproof cannot be used within the larger proof. The one exception to this is the single piece of information that the subproof "exports" back to the main proof. This single piece of information is that reported on the Show line. Once the subproof is completed and the Show line is cancelled, the subproof has successfully reported back to the main proof, and its output can then be used in the main proof. It is for this reason that line 7 of the above proof appeals to line 5 and 3, rather than 6 and 4. Lines 6 and 4, while they contain the Q and P that we want to form Q & P, are each "sealed away" inside a subproof; the only way we can get to that information is by having the subproofs report back via cancellation of their triggering Show lines.

Subproofs are thus like subroutines in a computer program. Typically, variable values assigned within a subroutine must be contained within that subroutine -- all the subroutine is allowed ultimately to do is to return some value (the value of the Show line). Information can, however, be moved downward from the main proof into the subproof. We see this happening on lines 4 and 6 where, while inside a subproof, we appeal to information in the main proof. Thus we have the fundamental principle of information flow in a proof structure:
Information can be transmitted downward (from a proof to a subproof) but not upward (from a subproof to a proof) except via successful cancellation of a Show line.

**Question #5**: How do I know when I can cancel a Show line?

**Answer**: At first glance, the answer to this question is obvious: when you've derived the sentence that the Show line commits you to deriving. In fact, the situation is a bit more complex. Suppose you set out on your proof with some Show line of the form \( \text{Show } \Phi \). Somewhere during this proof, you begin some subproof by invoking another Show line of the form \( \text{Show } \Theta \). During this subproof, you derive \( \Phi \). Can you then cancel the original Show line \( \text{Show } \Phi \)? No. To do so would be to export information upward from a subproof to a main proof, which is not allowed. So we get the following important principle:

Cancelling a Show line requires deriving the needed sentence in the proof triggered by that Show line, not in any subproof.

These principles about subproofs and information transmission between proofs and subproofs may well be rather obscure right now. Don't worry about it too much -- soon we'll introduce another kind of subproof and look at several examples that will help make the situation clearer.

**The Third Rule: Conditional Elimination** [Next]

At last, we are ready to go on to the next rule. Having set out elimination and introduction rules for the conjunction, it's time to move on to a new connective. We'll do the conditional next (largely because its rules are both useful and fairly easy to explain), starting with its elimination rule. Having rules for two different connectives will open up the proof system considerably, since we'll now be able to look at interactions between the two connectives, and we'll be able to prove some more interesting results.

The rule of conditional elimination, or \( \rightarrow E \), implements one of the most basic principles of reasoning. This principle tells us that if we have a conditional statement, such as "If it's raining, then the ground will get wet", and we have the antecedent of that conditional -- "It's raining" -- we can infer the consequence of the conditional ("The ground will get wet"). One way of thinking about this is that it's the point of the conditional ("if...then") to express potential inferences, so when we have a conditional and the first half of it, we are set to perform an inference to the second half. This type of reasoning has historically been called modus ponens (or, if you want to impress your friends, modus ponendo ponens).

Conditional elimination thus has the following input-output pattern:

\[
\rightarrow E:\quad \Phi \rightarrow \Theta \\
---------- \\
\Theta
\]

\( \rightarrow E \), like &I, requires two inputs. Unlike &I, however, the two inputs are not of equal status. One of the two inputs must be a conditional (that is, must have a conditional as its main connective), and the other must be the antecedent of that conditional. Given these two inputs, the consequent of the conditional can be derived using conditional elimination.

Let's look at a simple example using the rule of \( \rightarrow E \):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( P \land Q )</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>( Q \rightarrow R )</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Show ( R )</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( Q )</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>5</td>
<td>( R )</td>
<td>( \rightarrow E, 2,4 )</td>
</tr>
</tbody>
</table>

In this proof we have a conditional -- \( Q \rightarrow R \) -- which has as its consequent the sentence we want to prove (and that we set ourselves the goal of proving through the Show line). So to complete the proof, all we need to do is derive the antecedent of the conditional. We do this by applying &E to the conjunction \( P \land Q \), and extracting \( Q \). Then we are set to use \( \rightarrow E \) and complete the proof.

Some questions about the rule of \( \rightarrow E \):

**Question #1**: Can I use \( \rightarrow E \) to derive \( Q \) from \( P \rightarrow Q \) like this:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( P \land Q )</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>( Q \rightarrow R )</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Show ( R )</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( Q )</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>5</td>
<td>( R )</td>
<td>( \rightarrow E, 2,4 )</td>
</tr>
</tbody>
</table>

In this proof we have a conditional -- \( Q \rightarrow R \) -- which has as its consequent the sentence we want to prove (and that we set ourselves the goal of proving through the Show line). So to complete the proof, all we need to do is derive the antecedent of the conditional. We do this by applying &E to the conjunction \( P \land Q \), and extracting \( Q \). Then we are set to use \( \rightarrow E \) and complete the proof.
Answer: No. Definitely not. This is a common mistake that people make when they first start doing proofs. The rule of \( \rightarrow E \) requires two inputs -- both a conditional and the antecedent of the conditional. With just the conditional you cannot reach any conclusions at all.

**Question #2:** Why can't I use \( \rightarrow E \) in this way? That's the way that \&E worked.

**Answer:** True, but conjunctions are not the same as conditionals. If a conjunction \( \Phi \& \Theta \) is true, then -- by the semantic rule for \& -- both of its conjuncts must be true, so whichever conjunct you infer, you are guaranteed to be inferring something true. But when a conditional \( \Phi \rightarrow \Theta \) is true, all you know is that either \( \Theta \) is true or \( \Phi \) is false. You cannot safely conclude that \( \Theta \), because if \( \Phi \) is false, \( \Theta \) can also be false and the conditional true. Only once you know that \( \Phi \) (thus ruling out the possibility of making the conditional true by making \( \Phi \) false) can you safely conclude that \( \Theta \).

**Question #3:** The rule of \&E gave me a choice of two conclusions (either conjunct). Does \( \rightarrow E \) also give me a choice of two conclusions?

**Answer:** No. The only conclusion that \( \rightarrow E \) will ever license is the consequent of the conditional. In particular, you can never use \( \rightarrow E \) to derive the antecedent of the conditional.

**Question #4:** But what if I want to derive the antecedent of the conditional?

**Answer:** You frequently will want to derive it, because you will then use it in conjunction with \( \rightarrow E \) to derive the consequent of the conditional. But you can't get it using \( \rightarrow E \) (at least, not on that conditional). You'll have to find some other way. (We'll talk later about some general proof strategies that will help with this.)

**Question #5:** Does it matter whether the conditional or the antecedent of the conditional appears first in the proof?

**Answer:** No. They can be appealed to in either order. Both of the following are valid proofs:

1. \( P \rightarrow Q \) A
2. \( P \) A
3. \( \text{Show } Q \)
4. \( Q \rightarrow E, 1,2 \)

1. \( P \rightarrow Q \) A
2. \( P \rightarrow E, 1,2 \)
3. \( \text{Show } Q \)
4. \( Q \rightarrow E, 2,1 \)

**Question #6:** Does it matter in what order I cite the two lines that \( \rightarrow E \) requires as input?

**Answer:** No. I'll follow the convention of always listing the conditional first, but this isn't required.

**Question #7:** What if the antecedent or the consequent of the conditional are not sentence letters, but something more complicated? Can I still use \( \rightarrow E \)?

**Answer:** Yes. As long as you've got any sentence of the form \( \Phi \rightarrow \Theta \), and the corresponding sentence \( \Phi \) -- no matter how complex \( \Phi \) and \( \Theta \) are -- you can use \( \rightarrow E \). For example, all of the following are proper proofs:

1. \( (P \& Q) \rightarrow \neg(R \lor S) \) A
2. \( P \& Q \) A
3. \( \text{Show } \neg(R \lor S) \)
4. \( \neg(R \lor S) \rightarrow E, 1,2 \)

1. \( (P \rightarrow (R \leftrightarrow S)) \rightarrow \neg(Q \leftrightarrow (S \& \neg T)) \) A
2. \( P \rightarrow (R \leftrightarrow S) \) A
3. \( \text{Show } \neg(Q \leftrightarrow (S \& \neg T)) \)
4. \( \neg(Q \leftrightarrow (S \& \neg T)) \rightarrow E, 1,2 \)
(1) \((P \rightarrow Q) \rightarrow (R \rightarrow (S \rightarrow T))\) → \((S \rightarrow (S \rightarrow T))\)

(2) \((P \rightarrow Q) \rightarrow (R \rightarrow (S \rightarrow T))\)

(3) Show \(S \rightarrow (S \rightarrow T)\)

(4) \(S \rightarrow (S \rightarrow T)\) →E, 1, 2

Question #8: What if there's a conditional in the scope of another connective? Can I use →E to derive its consequent?

Answer: No. →E, like all the proof rules, can only be applied when its proper connective (here, the conditional) is the main connective of the input sentence. The following proofs are not correct:

(1) \((P \rightarrow Q) \lor (R \rightarrow S)\)

(2) \(P\)

(3) Show \(Q\)

(4) \(Q\) →E, 1, 2

(1) \(\neg (P \rightarrow (R \& S))\)

(2) \(P\)

(3) Show \(R \& S\)

(4) \(R \& S\) →E, 1, 2

(1) \(P \leftrightarrow (R \& (S \rightarrow T))\)

(2) \(S\)

(3) Show \(T\)

(4) \(T\) →E, 1, 2

In each of these "proofs", the first line does not have a conditional as its main connective, so the rule of →E cannot be applied, to derive these results or any other. Comparing these erroneous proofs with truth tables will quickly show that there is no implication in these cases. For example, in the first case take the interpretation:

- \(P = T\)
- \(Q = F\)
- \(R = F\)
- \(S = F\)

This interpretation makes true \(R \rightarrow S\), and hence \((P \rightarrow Q) \lor (R \rightarrow S)\) by way of verifying the second disjunct, but does not make true \(Q\), the conclusion of the argument.

Question #9: OK, what if I've got a conditional in the scope of another conditional? Then can I use →E on the smaller conditional, like this:

(1) \(P \rightarrow (Q \rightarrow R)\)

(2) \(Q\)

(3) Show \(R\)

(4) \(R\) →E, 1, 2

or maybe like this:

(1) \(P \rightarrow (Q \rightarrow R)\)

(2) \(Q\)

(3) Show \(P \rightarrow R\)

(4) \(P \rightarrow R\) →E, 1, 2

Answer: No. The rule →E must be applied to the main connective of the conditional. Both of the above proofs are incorrect. Interestingly, the second proof actually states a correct argument -- \(P \rightarrow (Q \rightarrow R)\), together with \(Q\), does indeed imply \(P \rightarrow R\) -- but the argument cannot be proved in this way (the correct proof will require a procedure of conditional introduction, which we'll come to next).

Question #10: But what if I've got a conjunction of two conditionals? Then can I use →E on one of the two, like this:

(1) \((P \rightarrow Q) \& (R \rightarrow S)\)
Answer: No. The conditional still is not the main connective, so you still cannot use $\rightarrow E$. However, it's easy to accomplish what you want to accomplish here. What we need to do is to isolate the conditional $R \rightarrow S$, so that the conditional will be the main connective. We can do this using &E:

| (1) $P \rightarrow Q$ & $(R \rightarrow S)$ | A |
| (2) R | A |
| (3) Show $S$ | |
| (4) $R \rightarrow S$ &E, 1 | |
| (5) $S$ $\rightarrow E$, 4, 2 | |

Question #11: What if I've got a conditional with a conjunction (disjunction) in the consequent of the conditional? Can I conclude just one of the conjuncts (disjuncts) like this:

| (1) $P \rightarrow (Q \& R)$ | A |
| (2) $P$ | A |
| (3) Show $Q$ | |
| (4) $Q$ $\rightarrow E$, 1, 2 | |

Answer: No. You could derive $Q \& R$ from the conditional using $\rightarrow E$, and then derive $Q$ using &E (although this obviously wouldn't work if the consequent contained a disjunction), but you can't go directly to $Q$. To use $\rightarrow E$, you must have one line which is a conditional (has a conditional as its main connective) and have another line which is just the antecedent of that conditional (and is the entirety of that antecedent), and then you can derive only the exact consequent of that conditional. There's no deviation from this in the slightest.

OK, enough questions. Before we look at a few more examples of applying $\rightarrow E$, I want to indulge in....

A Brief Digression (Skip)
The rule of $\rightarrow E$ often seems like the most undeniably valid rule of reasoning imaginable. What could be more certain than that I can reason from $If P then Q$ and $P$ to $Q$? Isn't that just what reasoning is -- combining facts with principles telling you that those facts lead to a certain conclusion to in fact draw that conclusion? However, an interesting challenge to the legitimacy of $\rightarrow E$ has recently been posed by the philosopher Vann McGee.

Here's an example of the problem that McGee points out (it's not the exact example he uses, but it has the same structure). Suppose three movies are up for the Best Picture Oscar:

- A Fistful of Pesos (a western)
- The Researchers (a western)
- Four Bar Mitzvahs and a Briss (a romantic comedy)

Given this information, we can obviously conclude that:

- If a western wins Best Picture, then if The Researchers doesn't win, A Fistful of Pesos will win.

Suppose further that the critical consensus is that The Researchers is the odds-on favorite (give it, say, a 75% chance of winning), that Four Bar Mitzvahs and a Briss is the next most likely (give it a 24% chance of winning), and that A Fistful of Pesos is a very dark horse (with a 1% chance of winning). On the basis of this, it seems reasonable to say that we are rationally warranted in concluding that:

- A western will win Best Picture.

(Of course, we don't know this with absolute certainty, but we pretty much never know anything with absolute certainty. 75% probability is as high as many things that we commonly say that we know.)

But now the problem emerges. We have seen that we are rationally warranted in the following two conclusions:

- If a western wins Best Picture, then if The Researchers doesn't win, A Fistful of Pesos will win.
- A western will win Best Picture.

Letting $P$ be A western will win Best Picture, and letting $Q$ be If The Researchers doesn't win, A Fistful of Pesos will win, we see that our two conclusions are of the form:

- $P \rightarrow Q$
- $P$

so, using $\rightarrow E$ we are warranted in concluding $Q$, or:

- If The Researchers doesn't win, A Fistful of Pesos will win.
But this conclusion seems clearly false. Given the critical consensus that I set out before, it looks like the right conclusion is:

- If The Researchers doesn't win, Four Bar Mitzvahs and a Briss will win.

After all, Four Bar Mitzvahs is the next strongest film -- why would it be rational to conclude that if the odds-on favorite fails, then the dark horse will win?

If our intuitions are right, and the conclusion:

- If The Researchers doesn't win, A Fistful of Pesos will win.

is in fact false, then this is a serious problem, because that conclusion is quite straightforwardly licensed by \( \rightarrow E \). I'm not going to try to solve this issue here (it's quite a complex one, involving a lot of delicate issues, many of which we'll touch on later, and there is as yet no general agreement about the right thing to say about the problem), but it's a good exercise to think about it. Keep in mind that very smart people who've thought a great deal about these issues disagree on what to say about the problem, so if you think you've got a simple solution to it, you're probably wrong.

End of Digression

We're just going to ignore this potential problem with \( \rightarrow E \) and push on. So, let's look at a some more elaborate proofs involving \( \rightarrow E \). We'll start with a proof of the following claim:

- \( P \land Q, Q \land R, (P \land R) \rightarrow S \), \( \therefore S \)

The proof proceeds as follows:

1. \( P \land Q \), A
2. \( Q \land R \), A
3. \( (P \land R) \rightarrow S \), A
4. \( \text{Show } S \)
5. \( P \rightarrow E, 1 \)
6. \( Q \rightarrow E, 2 \)
7. \( R \land S \rightarrow E, 3,7 \)
8. \( S \rightarrow E, 4,10 \)

Notice that here we have to disassemble the two conjunctions using elimination rules, and then reassemble the pieces to be able to use the conditional elimination to arrive at our goal. Finding the right strategy for this proof (and for many proofs) can be done by working backward -- starting with what we want as our final result (as announced in the Show line) and figuring out what rules applied to what will get us there. Knowing that we want to prove \( S \), we can see from the beginning -- due to the conditional on line 3 -- that if we could just get \( P \land R \) we could reach our goal. The puzzle thus becomes how to get \( P \land R \). Since the \&I rule is the obvious way to arrive at a conjunction, we have to think about what we need to use \&I to get \( P \land R \). The answer, of course, is that we need \( P \) and we need \( R \). Thus our goals shifts to deriving each of these. But these can each be derived through an application of \&E to one of our starting conjunctions.

Next let's look at an example that requires three applications of \( \rightarrow E \). We will prove the following claim:

- \( P \land Q, P \rightarrow R, Q \rightarrow S, (R \land S) \rightarrow T \), \( \therefore T \)

The proof proceeds as follows:

1. \( P \land Q \), A
2. \( P \rightarrow R \), A
3. \( Q \rightarrow S \), A
4. \( (R \land S) \rightarrow T \), A
5. \( \text{Show } T \)
6. \( P \rightarrow E, 1 \)
7. \( R \rightarrow E, 2,6 \)
8. \( Q \rightarrow E, 3,8 \)
9. \( S \rightarrow E, 4,7 \)
10. \( R \land S \rightarrow E, 7,9 \)
11. \( T \rightarrow E, 4,10 \)

Again, thinking backward can clarify the strategy of this proof. A step at a time, we see that:

- Given the conditional \( (R \land S) \rightarrow T \), we could get our goal \( T \) using \( \rightarrow E \) if we had \( R \land S \). Thus \( R \land S \) becomes our new goal.
- Our new goal \( R \land S \) could be obtained using \&I if we had \( R \) and we had \( S \).
- \( R \) could be obtained from \( P \rightarrow R \) using \&I, and \( S \) could be obtained from \( Q \rightarrow S \) using \( \rightarrow E \), if we had \( P \) and we had \( Q \).
- Finally, \( P \land Q \) can be immediately obtained from \( P \land Q \) using \&E.

Once we lay out this strategy, we just implement it -- in reverse order -- to create the proof.
Let's look at one more example, this time one involving nested conditionals. We will prove the following claim:

- $P \to (R \land ((R \land P) \to (S \to T))), P \land S, \therefore T$

The proof proceeds as follows:

1. $P \to (R \land ((R \land P) \to (S \to T))) \quad \text{A}\$
2. $P \land S \quad \text{A}\$
3. Show $T$
4. $P \land E, 2$
5. $R \land ((R \land P) \to (S \to T)) \quad \to E, 1, 4$
6. $R \quad \land E, 5$
7. $(R \land P) \to (S \to T) \quad \land E, 5$
8. $R \land P \quad \land I, 6, 4$
9. $S \to T \quad \to E, 7, 8$
10. $S \quad \land E, 2$
11. $T \quad \to E, 9, 10$

Although I won't go through the work of laying it out, the same sort of working-backward-from-the-end strategy will map a course through this proof as well.

The Fourth Rule: Conditional Introduction

We are finally ready to move on to the next rule -- the rule of conditional introduction. This rule will force us to introduce a new wrinkle or two into our proof system. To see why this is, let's start by trying to set up a standard rule of conditional introduction and see what goes wrong.

What we want is some sort of input-output filter which will convert some as-yet-unspecified inputs into a conditional of the form $\Phi \to \Theta$. This filter needs to meet two important conditions. First, it needs to be safe -- whenever it passes some inputs to the output $\Phi \to \Theta$, those inputs really need to guarantee the truth of that conditional (otherwise, our proof system will give bad results). Second, it needs to be adequately powerful -- it must let us derive all the conditionals we want to be able to derive.

We can get a start on the first condition by looking at the truth table for a conditional:

<table>
<thead>
<tr>
<th>$\Phi$</th>
<th>$\Theta$</th>
<th>$\Phi \to \Theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

By inspecting this truth table, we can see that the truth of the consequent guarantees the truth of the conditional, and that the falsity of the antecedents guarantees the truth of the conditional. This immediately suggests two possible rules of conditional introduction:

- $\Phi \quad \therefore \quad \Theta \to \Phi$

and:

- $\neg \Phi \quad \therefore \quad \Phi \to \Theta$

We could even allow both of these rules.

The problem, however, is that these rules don't allow us to derive many of the kinds of conditionals we want to be able to derive. Consider a typical conditional we might want to derive, such as:

- If this polynomial has degree three, then it has at most three distinct roots.

The two rules above would allow us to derive this conditional if we knew that the polynomial in question did not degree three, and it would also allow us to derive it if we knew that it had three roots. But there's a natural sense that we can know the truth of this conditional without knowing either of these things (in fact, in a moment I'll sketch a proof of how we...
would come to know this) -- that if we can know the conditional only by knowing exactly how things stand with respect to either the degree or the roots of the polynomial, then we've missed the very conditionality of the sentence in question, the way in which it expresses a link between two possible facts without committing to whether either of them holds.

The puzzle, then, is how we can craft a safe rule of conditional introduction which respects the conditionality of conditional claims (and thus doesn't require us to know definite states of the world to derive them). To help with solving this puzzle, let's look at how we might (informally, in English rather than in a formal logic) reason our way to the conditional given above. The following argument might be given:

Suppose that the polynomial is of degree three (although, of course, we don't really know for sure). Then it can be factored into some form \((ax - b)(cx - d)(ex - f)\). The roots of the equation will then be those \(x\) such that:
- \((ax - b)(cx - d)(ex - f) = 0\)
But, of course, a product of three numbers -- namely, \(ax - b, cx - d,\) and \(ex - f\) -- will be zero if and only if one of the three is zero. So we can have a root only if one of those three linear equations has a root. But linear equations can have only a single root, so the polynomial (continuing to assume it is of degree 3) must have at most three roots. So, if the polynomial is of degree three, it has at most three distinct roots.

Notice how this argument establishes the truth of the conditional. We begin by assuming -- with no justification -- that the antecedent of the conditional is true. We then reason a bit, until we deduce the consequence of the conditional, and then we assume that the conditional itself is true.

On the face of it, this should be an invalid form of reasoning. Why shouldn't the fact that we are helping ourselves to an unjustified assumption undermine the line of reasoning? But in fact this isn't a problem, for the following reason. The extra assumption we've made (without justification) must be either true or false (although we don't know which). Consider each case:
- If the extra assumption is true, then the consequent is true (since we've shown that it follows from the extra assumption, and that which follows from the true is itself true), and the truth of the consequent guarantees the truth of the whole conditional.
- On the other hand, if the extra assumption is false, then the antecedent of the conditional is false (since that is what the extra assumption is), so the conditional as a whole is true (since the falsity of the antecedent guarantees the truth of the conditional).

Even though we don't know whether the extra assumption is true or false, we do know that, either way, the conditional we wanted is true. Thus making the extra assumption allows us to prove the conditional.

We are now ready to give the technique for introducing conditionals in our proof system. Suppose we want to prove some conditional of the form \(\Phi \rightarrow \Theta\). We take the following steps:
1. Write a Show line declaring an intention to derive \(\Phi \rightarrow \Theta\). Doing so triggers the beginning of a subproof.
2. Introduce the antecedent \(\Phi\) of the conditional as a new assumption.
3. Within the subproof, attempt to derive the consequent \(\Theta\) using our standard proof rules.
4. Once \(\Theta\) has been derived, end the subproof and cancel the Show line from step 1.

Let's look at a very simple example to see how this works. We'll derive the following claim:
- \(P \rightarrow (Q \rightarrow R), Q, \therefore P \rightarrow R\)

The proof proceeds as follows:

| (1) \(P \rightarrow (Q \rightarrow R)\) | A |
| (2) \(Q\) | A |
| (3) Show \(P \rightarrow R\) | |
| (4) \(P\) | ACP |
| (5) \(Q \rightarrow R\) | \(\rightarrow E, 1,4\) |
| (6) \(R\) | \(\rightarrow E, 2,5\) |

ACP, in the justification of line 4, means Assumption for Conditional Proof, and is the justification we'll use for the extra assumption we use in doing conditional introduction.

This proof has significant differences from the proofs we've done thus far, so let's start by going over some questions about it.

Question #1: So where's the rule of conditional introduction? I don't see \(\rightarrow I\) used anywhere.

Answer: There is no rule of conditional introduction. The proof above derives the conditional not by invoking a particular proof rule, but by employing a particular proof strategy. Proof rules attached to regular lines, and explain how the content
of that line can be derived from the given input lines. Proof strategies, on the other hand, are associated with Show lines, and specify a method for achieving the goal announced in the Show line.

Prior to considering conditional introduction, we had only a single proof strategy, so the very fact that there was such a thing as a proof strategy never came up. Our initial, pre-conditional-introduction, proof strategy was what we call the **strategy of direct proof**.

**Direct Proof**: In direct proof, a Show line declares an intention to show some claim. That claim can be anything the prover desires. In order to complete the triggered direct subproof, some later line of the proof must contain the claim the Show line declares an intention to show.

We've now introduced the **new strategy of conditional proof**:

**Conditional Proof**: In conditional proof, a Show line declares an intention to show a claim whose main connective is a conditional. The triggered conditional subproof then begins by adding as a new assumption (using ACP) the antecedent of the conditional to be proved. The subproof is then completed (and the Show line cancelled) when the consequent of the conditional has been derived.

Every time we write a Show line in a proof, we must invoke some proof strategy (thus far, either direct or conditional proof, although we will add more proof strategies as we go). Let's look at another example. We'll prove the following claim:

\[ P \rightarrow Q, Q \rightarrow R, \therefore P \rightarrow R \]

This is, in essence, the claim that implication is transitive, and that multiple proofs (that from \( P \) to \( Q \) and that from \( Q \) to \( R \)) can be stringed together to form a larger proof (from \( P \) to \( R \)). The proof proceeds as follows:

| (1) \( P \rightarrow Q \) | A |
| (2) \( Q \rightarrow R \) | A |
| (3) Show \( P \rightarrow R \) | [Conditional Proof] |
| (4) \( P \) | ACP |
| (5) \( Q \rightarrow E, 1,4 \) |
| (6) \( R \rightarrow E, 2,5 \) |

(Here I've added a notation after the Show line indicating which proof strategy we're invoking. I won't in general do this, but if it's helpful for you, you should include it.) Since we are invoking the strategy of conditional proof, we get to add the additional assumption (by ACP) that \( P \), and our success condition is not \( P \rightarrow R \) (as it would be if we were to proceed by direct proof), but rather \( R \) itself.

**Question #2**: In both of these proofs using the strategy of conditional proof, I don't actually see the conditional you wanted to prove showing up anywhere. Where is it?

**Answer**: When fulfilling a Show line invoked under the strategy of conditional proof (unlike one invoked under the strategy of direct proof), the success conditions do not require producing the sentence given on the Show line. Thus the conditional does not in fact appear anywhere other than on the original Show line. Once you've derived the consequent (as on line 6 of the previous proof), you have finished the conditional proof.

There wouldn't be anything wrong with a proof system which included an explicit move of conditional elimination, in which the conditional was formed once the consequent was derived. If we had such a system, then the previous proof would look like this:

| (1) \( P \rightarrow Q \) | A |
| (2) \( Q \rightarrow R \) | A |
| (3) Show \( P \rightarrow R \) | [Conditional Proof] |
| (4) \( P \) | ACP |
| (5) \( Q \rightarrow E, 1,4 \) |
| (6) \( R \rightarrow E, 2,5 \) |
| (7) \( P \rightarrow R \rightarrow I, 4,6 \) |

(I've put the proof in green here to show that, while not really wrong, it is non-standard, and not strictly in our proof system.)
Here \( P \rightarrow R \) is explicitly derived on line 7, and used as a basis for cancelling the \textit{Show} on line 3. If it's helpful for you to put in this extra step, you should feel free to do so. Again, there's nothing wrong with setting up the proof system this way -- it's just that the extra step isn't necessary, so for streamlining we omit it form the official system.

**Question #3:** Can a proof contain more than one conditional proof?

**Answer:** Yes. As an example, we'll prove the following claim:

- \( P \rightarrow (Q \& R) \), \( \therefore (P \rightarrow Q) \& (P \rightarrow R) \)

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( P \rightarrow (Q &amp; R) ) A</td>
</tr>
<tr>
<td>2</td>
<td>\textit{Show} ( (P \rightarrow Q) &amp; (P \rightarrow R) )</td>
</tr>
<tr>
<td>3</td>
<td>\textit{Show} ( P \rightarrow Q )</td>
</tr>
<tr>
<td>4</td>
<td>( P ) ACP</td>
</tr>
<tr>
<td>5</td>
<td>( Q &amp; R ) \rightarrow E, 1,4</td>
</tr>
<tr>
<td>6</td>
<td>( Q ) &amp;E, 5</td>
</tr>
<tr>
<td>7</td>
<td>\textit{Show} ( P \rightarrow R )</td>
</tr>
<tr>
<td>8</td>
<td>( P ) ACP</td>
</tr>
<tr>
<td>9</td>
<td>( Q &amp; R ) \rightarrow E, 1,8</td>
</tr>
<tr>
<td>10</td>
<td>( R ) &amp;E, 9</td>
</tr>
<tr>
<td>11</td>
<td>( (P \rightarrow Q) &amp; (P \rightarrow R) ) &amp;I, 3,7</td>
</tr>
</tbody>
</table>

In this proof, one conditional proof is triggered by the \textit{Show} on line 3. This conditional proof runs from lines 4 to 6, starting with the introduction of the antecedent of the conditional (that is, \( P \)) on line 4 using \textit{ACP}, and ending with the derivation of the consequent of the consequent of the conditional (that is, \( Q \)) on line 6. A second conditional proof is then triggered by the \textit{Show} on line 7. This second conditional proof runs from lines 8 to 10, again starting with the introduction of the antecedent, and ending with the derivation of the consequent. Once both conditionals have been derived, they are combined into the desired conjunction.

**Question #4:** Why do I have to introduce \( P \) again in line 8, and then derive the conjunction \( Q \& R \) again in line 9?

**Answer:** This is a result of the principle that information can only be transmitted downward from a larger proof to a subproof of that proof. Re-using \( P \), or \( Q \& R \), would require transmitting information horizontally, from one subproof to another (or, if you prefer, first upward to the main proof, and then downward to the second subproof -- either way, it is not allowed).

Of course, this just raises the question: why is it so important that information be transmitted only downward? To see why this is the case, suppose we lift for a moment the ban on transmitting information sideways (from one subproof to another). We could then prove the following:

- \( P \rightarrow (Q \& R) \), \( \therefore (P \rightarrow Q) \& (T \rightarrow R) \)

in the following manner:

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( P \rightarrow (Q &amp; R) ) A</td>
</tr>
<tr>
<td>2</td>
<td>\textit{Show} ( (P \rightarrow Q) &amp; (T \rightarrow R) )</td>
</tr>
<tr>
<td>3</td>
<td>\textit{Show} ( P \rightarrow Q )</td>
</tr>
<tr>
<td>4</td>
<td>( P ) ACP</td>
</tr>
<tr>
<td>5</td>
<td>( Q &amp; R ) \rightarrow E, 1,4</td>
</tr>
<tr>
<td>6</td>
<td>( Q ) &amp;E, 5</td>
</tr>
<tr>
<td>7</td>
<td>\textit{Show} ( T \rightarrow R )</td>
</tr>
<tr>
<td>8</td>
<td>( T ) ACP</td>
</tr>
<tr>
<td>9</td>
<td>( R ) &amp;E, 5</td>
</tr>
<tr>
<td>10</td>
<td>( (P \rightarrow Q) &amp; (T \rightarrow R) ) &amp;I, 3,7</td>
</tr>
</tbody>
</table>

Line 9 of this proof is illegitimate, because it invalidly imports information sideways from the first subproof, by taking line 5 (part of that first subproof) as an input to the \&E rule. And, importantly, the result that is derived here is incorrect. Suppose we have the following interpretation:

- \( P: F \)
- \( Q: F \)
- \( R: F \)
- \( T: T \)
This interpretation makes $P \rightarrow (Q \& R)$ true, by making its antecedent false. However, the interpretation also makes $T \rightarrow R$ false, by making its antecedent true and its consequent false. Thus the interpretation makes the premise true and the conclusion false, showing that this is not a valid argument. It’s to avoid being able to derive invalid results that we impose the ban on sideways transmission of information.

Upward transmission of information also causes problems, perhaps even more straightforwardly. Consider the claim:

- $P \rightarrow (Q \& R), \therefore Q$

This claim is clearly incorrect. From the conditional claim that if $P$ is true, $Q \& R$ is true, one cannot determine anything about the truth value of $Q$. It could be either true or false. However, the following incorrect "proof", using upward transmission of information, manages to "derive" this result:

| (1) $P \rightarrow (Q \& R)$ | A |
| (2) Show $Q$ | |
| (3) Show $P \rightarrow R$ | |
| (4) $P$ | ACP |
| (5) $Q \& R$ | $\rightarrow E, 1, 4$ |
| (6) $R$ | $\& E, 5$ |
| (7) $Q$ | $\& E, 4$ |

In this "proof", line 7 illegitimately transmits information upward from the subproof (running from lines 4 to 6) to the main proof.

In general, sideways and upward transmission of information is illegitimate because subproofs often have auxiliary hypotheses (such as those introduced by ACP) which are not really justified, but are just used as hypotheticals to determine the truth of a more complex statement (such as the conditional derived using the conditional subproof). If we export information upward from such a subproof, it will likely be infected by the unjustified auxiliary hypothesis, and thus will introduce unjustified assumptions into the higher-level proof.

**Question #5:** Back in the proof given in Question #3, why can the $\& I$ on line 10 appeal to the two Show lines of lines 3 and 7? I thought Show lines couldn't be used as inputs.

**Answer:** Each of these Show lines has been cancelled by the time line 10 is reached, because each of their subproofs has been successfully completed. Once a Show line has been cancelled, then the sentence in question has in fact been shown, so it is a legitimate input to other proof rules. Before it is cancelled, however, the Show line cannot be used as input to proof rules.

**Question #6:** Can I have one conditional subproof within another?

**Answer:** Yes. Subproofs can be nested as deeply as you want. Consider the proof of the following claim:

$P \rightarrow Q, \therefore P \rightarrow (R \rightarrow (S \rightarrow Q))$

The proof proceeds as follows:

| (1) $P \rightarrow Q$ | A |
| (2) Show $P \rightarrow (R \rightarrow (S \rightarrow Q))$ | |
| (3) $P$ | ACP |
| (4) Show $R \rightarrow (S \rightarrow Q)$ | |
| (5) $R$ | ACP |
| (6) Show $S \rightarrow Q$ | |
| (7) $S$ | ACP |
| (8) $Q$ | $\rightarrow E, 1, 3$ |

This proof starts off trying to show the iterated conditional $P \rightarrow (R \rightarrow (S \rightarrow Q))$. We thus take the antecedent -- $P$ -- of the conditional as an extra assumption using ACP, and attempt to prove the consequent $R \rightarrow (S \rightarrow Q)$. This consequent, however, is itself a conditional. Thus we invoke conditional proof again (starting a nested subproof), now taking the antecedent $R$ as an assumption and trying to prove $S \rightarrow Q$. Once again what we are trying to show is a conditional, so we invoke conditional proof a third time, now taking $S$ as an assumption and trying to prove $Q$. $Q$ is not a conditional, so we cannot invoke conditional proof again. Instead, we try to use proof rules to produce $Q$, which is straightforwardly done (by combining our starting assumption $P \rightarrow Q$ with our first ACP $P$ -- note that this requires downward transmission of information from the first subproof to the third (nested) subproof).
Question #7: In that previous proof, it doesn't seem like the ACP assumptions R (from line 5) and S (from line 7) are doing any work -- they never get used as input to any rule, do they?

Answer: No, they don't. In a sense, they aren't doing any work. In a way, what we are showing in the above proof is that if we have an argument from some premise to a conclusion (that is, we have P → Q), then adding new premises (that is, R and S to make P → (R → (S → Q))) doesn't undermine the validity of the argument. That's because the new premises can simply be inert -- as long as you've still got P, you've still got what you need to reach Q. So really it's no surprise that R and S do nothing in the proof -- what the proof is showing is that they don't have to do anything.

Question #8: What if the antecedent of the conditional I'm trying to prove isn't the most helpful assumption for me to make? Can I assume something else using ACP?

Answer: No. ACP can only be used to introduce the antecedent of a conditional you're trying to prove as the beginning of a conditional proof. Using it to introduce other claims will lead to false results. For example, consider the following fallacious "proof":

<table>
<thead>
<tr>
<th>Line</th>
<th>Premise/Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
</tr>
<tr>
<td>2</td>
<td>Show P → (R &amp; Q)</td>
</tr>
<tr>
<td>3</td>
<td>Q</td>
</tr>
<tr>
<td>4</td>
<td>R &amp; Q</td>
</tr>
</tbody>
</table>

This proof illegitimately introduces Q (rather than P) with ACP. As a result, it is able to prove a false result. An interpretation like this:
- P: T
- Q: F
- R: T

will make the initial assumption P true and the conclusion P → (Q & R) false. (Of course, that interpretation makes the auxiliary hypothesis Q false, but there is no requirement that auxiliary hypotheses come out true -- they are supposed to be inessential hypotheses).

Question #9: Can I use ACP at other times, when I'm not starting a conditional proof?

Answer: No. ACP can only ever be used as the first step in a conditional proof.

Question #10: Can I use the very same line to both begin and end the conditional proof?

Answer: Yes. This actually happens when proving a condition in which the antecedent and the consequent are the same thing. Consider how we would prove the following claim:
\[ \vdash \neg P \rightarrow \neg P \]
(Remember that an implication claim with no premises given is the same as the claim that the conclusion sentence is valid.) The proof would proceed as follows:

<table>
<thead>
<tr>
<th>Line</th>
<th>Premise/Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Show \neg P \rightarrow \neg P</td>
</tr>
<tr>
<td>2</td>
<td>\neg P</td>
</tr>
</tbody>
</table>

On line 2, we begin the conditional proof by assuming the antecedent of the conditional -- that is, \neg P. Our goal is now to derive, within the conditional proof, the consequent \neg P of the conditional. But, of course, we already have \neg P, so we don't have to do anything to achieve our goal. Thus the proof comes to a very swift end.

Question #11: Wait -- that last proof started with a Show line, and didn't have any initial assumptions using the A rule. Can you do that?

Answer: Yes. If you're trying to prove an implication claim in which no premises are given, then your proof won't have any initial assumptions, and will launch right into the Show line. This will only be possible if the Show line is triggering a proof type which grants an auxiliary hypothesis (like conditional proof) -- otherwise there will be no inputs for the proof rules to work off of.

Question #12: Do I have to be trying to show a conditional in order to use conditional proof?

Answer: Yes. The strategy of conditional proof can only be invoked when the sentence on the Show line is a conditional.
Question #13: What if there's a conditional somewhere in what I'm trying to show, but it's not the main connective? Then can I use conditional proof?

Answer: No. Just as the proof rules are sensitive to the main connectives of sentences, so are the proof strategies. Thus none of the following lines can be used to trigger a conditional proof:

• Show \((P \to Q) \lor R\)
• Show \(~(P \to R)\)
• Show \(P \leftrightarrow Q\)
• Show \((P \to (Q \to R)) \leftrightarrow S\)

Question #14: Why do you sometimes refer to conditional proofs and sometimes to conditional subproofs?

Answer: This is mostly a terminological issue. We call a proof any process of establishing the correctness of a Show line. All of the proofs that we give here have some primary Show line, and the (main) proof is the process of establishing the claim made on that Show line. However, frequently in the process of establishing this, we'll introduce additional Show lines. Each of these new Show lines gives rise to a new proof, but since these new proofs are components of a larger proof, then we call them subproofs, just to clarify their position in the overall structure.

For example, consider the following proof:

| (1) \((P \to P) \to (Q \& R)\) | A  |
| (2) Show \(Q \& R\) |
| (3) Show \(P \to P\) |
|   | (4) \(P\) | ACP |
| (5) \(Q \& R\) | \(\to\)E, 1,3 |

This proof (we will also use the word proof just to refer to the total sequence of lines) contains two proofs. One is the main proof (of \(Q \& R\), running the entire length of the proof, from lines 1 to 5. The other is the subproof (of \(P \to P\), on lines 3 and 4.

Question #15: Do I have to use conditional proof when I'm trying to prove a conditional?

Answer: No. Conditional proof can only be used when the sentence on the Show line is a conditional, but direct proof can be used at any time, including when trying to show a conditional. In most cases, conditional proof will work much better (because the extra ACP helps get the chain of inferences moving), but in some cases direct proof makes more sense. Thus consider the claim:

\((P \to Q) \& (R \to S) \therefore P \to Q\)

This claim is most easily proved using direct proof:

| (1) \((P \to Q) \& (R \to S)\) | A  |
| (2) Show \(P \to Q\) |
| (3) \(P \to Q\) | \&E, 1 |

(Notice that we do not take \(P\) using ACP here, since we are not invoking the strategy of conditional proof). However, it can also be proved using conditional proof:

| (1) \((P \to Q) \& (R \to S)\) | A  |
| (2) Show \(P \to Q\) |
| (3) \(P\) | ACP |
| (4) \(P \to Q\) | \&E, 1 |
| (5) \(Q\) | \(\to\)E, 4,3 |

The proof via conditional proof is, in this case, clearly derivative on and less efficient than the direct proof.

More Examples of Conditional Proof [Next]

Let's now do a few more examples using conditional proof, some a bit more complex than those we've done so far.

For our first example, we will prove:

• \(\therefore ((P \& Q) \to R) \to ((P \to (Q \to R))) \& ((P \to (Q \to R)) \to ((P \& Q) \to R))\)
This proof will, in essence, show the two forms \((P \& Q) \rightarrow R\) and \(P \rightarrow (Q \rightarrow R)\) are equivalent, by proving conditionals in each direction between the two. The proof proceeds as follows:

(1) Show \(((P \& Q) \rightarrow R) \rightarrow ((P \rightarrow (Q \rightarrow R)) \& ((P \rightarrow (Q \rightarrow R)) \rightarrow ((P \& Q) \rightarrow R)))\)

(2) Show \(((P \& Q) \rightarrow R) \rightarrow ((P \rightarrow (Q \rightarrow R)))\)

(3) \((P \& Q) \rightarrow R\) ACP

(4) Show \(P \rightarrow (Q \rightarrow R)\)

(5) \(P\) ACP

(6) Show \(Q \rightarrow R\)

(7) \(Q\) ACP

(8) \(P \& Q\) &I, 5, 7

(9) \(R\) \( \rightarrow\) E, 3, 8

(10) Show \(((P \rightarrow (Q \rightarrow R)) \rightarrow ((P \& Q) \rightarrow R))\)

(11) \(P \rightarrow (Q \rightarrow R)\) ACP

(12) Show \((P \& Q) \rightarrow R\)

(13) \(P \& Q\) ACP

(14) \(P\) &E, 13

(15) \(Q \rightarrow R\) \( \rightarrow\) E, 11, 14

(16) \(Q\) &E, 13

(17) \(R\) \( \rightarrow\) E, 15, 16

(18) \(((P \& Q) \rightarrow R) \rightarrow ((P \rightarrow (Q \rightarrow R))) \& ((P \rightarrow (Q \rightarrow R)) \rightarrow ((P \& Q) \rightarrow R))\) &I, 2, 10

On its top level of structure, this proof proceeds by direct proof (a direct proof which will culminate in line 18, as &I is used to paste together two conditionals). That direct proof then contains two subsidiary conditional proofs, which establish the conditionals in both directions between \((P \& Q) \rightarrow R\) and \(P \rightarrow (Q \rightarrow R)\). Each of these subproofs contains further conditional subproofs – nested two levels deep in the first case and one level deep in the second case. It's an excellent exercise to go through and make sure the rules on information transmission are correctly followed at every step in this proof. If you can thoroughly follow this proof, you're got a fully adequate grasp of the proof system thus far.

The result just proved is an important one because it shows that there is no significant difference between a single proof containing a number of premises (which is what \((P \& Q) \rightarrow R\) amounts to) and a series of proofs chained together (which is what \(P \rightarrow (Q \rightarrow R)\) amounts to). It's thus, in effect, an illustration of the monotonicity of deductive reasoning.

Next, we will prove the following result:

\[
\therefore (P \rightarrow Q) \rightarrow ((R \rightarrow (Q \rightarrow S)) \rightarrow (R \rightarrow (P \rightarrow S)))
\]

This result tells us, in essence, that we can replace a premise \((Q)\) in an extended proof with a logically stronger premise \((P\), logically stronger because we have \(P \rightarrow Q\)) without disturbing the validity of the proof. The proof proceeds as follows:

(1) Show \((P \rightarrow Q) \rightarrow ((R \rightarrow (Q \rightarrow S)) \rightarrow (R \rightarrow (P \rightarrow S)))\)

(2) \(P \rightarrow Q\) ACP

(3) Show \((R \rightarrow (Q \rightarrow S)) \rightarrow (R \rightarrow (P \rightarrow S))\)

(4) \(R \rightarrow (Q \rightarrow S)\) ACP

(5) Show \(R \rightarrow (P \rightarrow S)\)

(6) \(R\) ACP

(7) Show \(P \rightarrow S\)

(8) \(P\) ACP

(9) \(Q \rightarrow S\) \( \rightarrow\) E, 4, 6

(10) \(Q\) \( \rightarrow\) E, 2, 8

(11) \(Q\) \( \rightarrow\) E, 10, 11

For our last example, we will prove the following result:

\[
\therefore ((P \rightarrow Q) \rightarrow ((R \& U) \rightarrow S) \& V \& U, (Q \rightarrow R) \rightarrow (P \rightarrow S))
\]

The proof proceeds as follows:

(1) \((P \rightarrow Q) \rightarrow ((R \& U) \rightarrow S) \& T\) A

(2) \(V \& U\)

(3) Show \((Q \rightarrow R) \rightarrow (P \rightarrow S)\)

(4) \(Q \rightarrow R\) ACP

(5) Show \(P \rightarrow S\)
The Fifth Rule: Negation Elimination [Next]

We're now ready to move on to the rules governing negation. As before, there will be an elimination and an introduction procedure for negation, although (as with the conditional) there won't strictly speaking be a rule for each.

We'll start with the rule of negation elimination. This rule isn't going to do quite what one might expect it to do. Given the earlier examples of &E and →E, one would expect a rule which takes an input of the form ¬Φ, perhaps in combination with some other input, and produced an output without the negation -- preferably, an output of the form Φ. However, this is not what we will get. Instead, we will get a rule which eliminates two negations at once. For this reason, we'll call the rule Double Negation, or ¬¬, instead of Negation Elimination as one might expect.

The input-output conditions for ¬¬ are as follows:


Notice that instead of the usual single line between the input requirements and the output results, I've put a double line this time. That's because ¬¬ is a bidirectional rule. You can either take ¬¬Φ as input, get rid of the double negation, and obtain Φ as output, or you can take Φ as input, add a double negation, and obtain ¬¬Φ as output. Both are permissible. ¬¬ is thus really a combination of two monodirectional rules:

and:

Both of these inferential moves are sensible ones. Given that the semantic effect of negation is to reverse truth values (change true to false and false to true), a double negation has the effect of a double reversal, which leaves everything back the way it was. Double-negating a sentence is like rotating it 180º twice, and thus travelling a full circle. (Were it not for typographical obstacles, we could actually have used inversion as our method marking negation, which would have eliminated the need for a rule of double negation.)

Double negation is an extremely straightforward rule to apply, so there's not a great deal to say about it. Let's look at a few examples. First, we'll use the rule in the elimination direction to prove the following:

• P & ¬¬Q, (Q & P) → ¬¬R, ∴ R

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>(1) P &amp; ¬¬Q A</th>
<th>(2) (Q &amp; P) → ¬¬R A</th>
<th>(2) Show R</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3) P &amp;E, 1</td>
<td>(4) ¬¬Q &amp;E, 1</td>
<td>(5) Q ¬¬, 4</td>
</tr>
<tr>
<td>(6) Q &amp; P &amp;I, 5,3</td>
<td>(7) ¬¬R →E, 2,6</td>
<td></td>
</tr>
</tbody>
</table>
Next, we'll use the rule in the introduction direction to prove the following:

• \( P \rightarrow Q, \therefore P \rightarrow \neg
\neg Q \)

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>(1) ( P \rightarrow Q )</th>
<th>A</th>
</tr>
</thead>
</table>
| (2) Show \( P \rightarrow \neg
\neg Q \) | ACP |
| (3) \( P \) | \( \neg
\neg E, 1,3 \) |
| (4) \( Q \) | \( \neg
\neg 4 \) |
| (5) \( \neg
\neg Q \) | \( \neg
\neg 5 \) |

Finally, we'll use the rule in both directions to prove the following:

• \( \neg
\neg P \rightarrow \neg
\neg (Q \& \neg
\neg R), \therefore P \rightarrow R \)

The proof proceeds as follows:

| (1) \( \neg
\neg P \rightarrow \neg
\neg (Q \& \neg
\neg R) \) | A |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Show ( P \rightarrow R )</td>
<td>ACP</td>
</tr>
</tbody>
</table>
| (3) \( P \) | \( \neg
\neg 3 \) |
| (4) \( \neg
\neg (Q \& \neg
\neg R) \) | \( \neg
\neg E, 1,4 \) |
| (5) \( Q \& \neg
\neg R \) | \( \neg
\neg 5 \) |
| (6) \( R \) | \( \&E, 6 \) |
| (7) \( \neg
\neg R \) | \( \neg
\neg 7 \) |

A few quick questions about the \( \neg
\neg \) rule:

**Question #1:** Can I use \( \neg
\neg \) to eliminate/introduce of more (or fewer) than two negations?

**Answer:** No. Only **double** negations can be introduced or eliminated using the \( \neg
\neg \) rule. If I have a sentence with a single negation, I cannot eliminate it -- although I can add two new negations to form a triple negation. If I have a triple negation, I can remove two of those negations to form a single negation. If I have a quadruple negation, I can remove two negations to form a double negation, and then use the rule again (on another line) on that double negation to form an unnegated sentence -- but I cannot remove all four negations at once. The effect is thus that, through enough applications of the \( \neg
\neg \) rule, I can completely eliminate any even-numbered string of negations, and convert any odd-numbered string of negations into a single negation.

**Question #2:** Couldn’t we just add a general rule allowing you to get rid of any even number of negations?

**Answer:** Yes, we could (although a full justification of this rule on the basis of the \( \neg
\neg \) rule would force us to use some techniques of mathematical induction that we haven't introduced yet). But there's really not much point in cluttering up the proof system with this, since odds are we'll never actually encounter a string of four or more negations. I'll make you a deal -- if you're sure that you understand the rule of \( \neg
\neg \), and why you strictly can’t use it to get rid of four (or six, or etc.) negations at once, you’ve got my permission to go ahead and get rid of any even number of negations in one fell swoop, should the opportunity ever arise.

**Question #3:** Can I use \( \neg
\neg \) to get rid of an imbedded double negation like this:

| (1) \( P \rightarrow \neg
\neg Q \) | A |
|----------------------|---|
| (2) Show \( P \rightarrow Q \) | \( \neg
\neg 1 \) |

**Answer:** No. Like all the other rules, \( \neg
\neg \) requires the relevant connective to be the main connective of the sentence. However, in the case of \( \neg
\neg \), we'll qualify this position somewhat when we discuss the procedure of Replacement below.

**The Sixth Rule: Negation Introduction [Next]**

The rule of \( \neg
\neg \) only allows us to a small amount of what we want to do, inferentially speaking, with negation. What we need now is a way of introducing (and, as we'll show, also eliminating) single negations. As with conditional elimination, it's going to turn out that what we want here is a proof strategy, rather than a specific proof rule.
Let's think a bit, then, about what we want to do and how we can do it. We're looking for some way to introduce a single negation -- some way to reach conclusions of the form \( \neg \Phi \). It turns out that it's more helpful here to think about trying to prove \( \Phi \) false, rather than trying to prove \( \neg \Phi \) true (notice that these are trivially equivalent results). Suppose, then, that there is a claim you want to show false. How does one go about doing this?

Let's set aside the formalism of the logical system for a moment and think about some everyday examples. Sometimes one is in the fortunate position of being able directly to prove that a claim is false. Suppose, for example, that someone asserts You only have four fingers on your left hand. I can prove this false simply by holding up my left hand and displaying its five fingers. Here I am in essence proving the claim false by just introducing its negation as true (on the basis of direct evidence from my senses). Most of the time, however, one is not so fortunate as this. Suppose, then, that a friend asserts John F. Kennedy was actually the same person as John Lennon, and you want to prove this false. Here you can't simply hold up evidence to the distinctness of Kennedy and Lennon.

So what do you do? One way to proceed is to start point out that various bizarre, absurd, or otherwise unacceptable consequences follow from your friend's assertion. You might, for example, point out that it follows from what your friend says that either Kennedy wasn't really killed in Dallas in 1963, or Kennedy/(Lennon) was in fact killed twice. Or you might point out that it follows that the president of the United States was a British citizen who was 20 years old when he was elected president. Or you might point out that it follows that Kennedy was recording "Love Me Do" during the Cuban missile crisis. Since (so goes your argument) all of these things are absurd, it follows that the original claim (from which they follow) that Kennedy was the same person as Lennon must be false.

In arguing in this way, you are implicitly appealing to the nature of valid arguments. Valid arguments take you from true premises to a true conclusion. So if you can present a valid argument (If Kennedy was Lennon, then Kennedy was recording "Love Me Do" during the Cuban missile crisis), and show that the conclusion of that argument is false (by observing the obvious absurdity of that conclusion), you can safely conclude that the premise is also false (and hence that John F. Kennedy was actually the same person as John Lennon, and you want to prove this false).

The above example was one of informal reasoning, so the notions of "following from" and "absurd consequence" were both rather loose. We can see a somewhat more strict application of the same general form of reasoning by considering the proof (due to the classical Greek mathematical school of Pythagoras) that the square root of two is not a rational number:

Suppose that the square root of two were some rational number -- in particular, that it were a maximally reduced fraction \( n/m \) (where \( n \) and \( m \) have no common factors). Then it would follow that \( (n/m)^2 = n^2/m^2 = 2 \), and thus that \( 2m^2 = n^2 \). Since \( n^2 \) equals a multiple of 2, \( n^2 \) must be even. Since the square of an odd number is odd, it follows that \( n \) itself must be even, so we have \( n = 2k \), for some \( k \). Thus we know that \( (2k/m)^2 = 4k^2/m^2 = 2 \), and thus that \( 2m^2 = 4k^2 \). From this it follows that \( m^2 = 2k^2 \), and hence that \( m^2 \) is even. By the same reasoning as before, \( m \) must be even, so \( m = 2j \), for some \( j \). But now we have \( n/m = 2k/2j \), which contradicts the assumption that \( n/m \) is a maximally reduced fraction. Thus our assumption was false, and the square root of two is not a rational number.

This proof displays the level of rigour that we're going to require of our proof system, so we're now ready to set out the procedure for proving a negative. We will introduce a new proof strategy called the strategy of Indirect Proof. Since it's a proof strategy, it will be invoked through the introduction of a Show line, and will specify how the triggered subproof is carried out. Indirect proof is characterized as follows:

- The strategy of indirect proof can be invoked only in response to a Show line in which the sentence to be shown has a negation as its main connective.
- In an indirect proof, one then adds as extra assumption the unnegated form of the sentence on the Show line, giving as justification AIP (Assumption for Indirect Proof).
- The subproof then has the goal of producing an explicit contradiction, which must take the form of two lines, one of which is the negation of the other.
- Once the explicit contradiction has been derived, the subproof ends and the Show is cancelled.
The indirect proof thus proves a claim false (i.e., proves its negation true) by temporarily (and hypothetically) assuming that claim to be true and showing that its truth would lead to a contradiction, and hence must be rejected.

Let's look at a couple of examples. First we'll look at a fairly minimalist application of the strategy of indirect proof, by proving:

- \( P \rightarrow \neg P \)

The proof proceeds as follows:

| (1) \( P \rightarrow \neg P \) | A |
| (2) Show \( \neg P \) | |
| (3) \( P \) | AIP |
| (4) \( \neg P \) | \( \rightarrow E, 1, 2 \) |

In response to the Show on line 2, we invoke the strategy of indirect proof. This is permissible because what we are trying to prove -- \( \neg P \) -- has a negation as its main connective. Indirect proof then tells us to take the unnegated form of what we're trying to prove -- hence, \( P \) -- and add it as an auxiliary assumption using AIP. Our goal is now to reach an explicit contradiction. Given that we've already got \( P \) in the subproof, one easy way of reaching such a contradiction would be by deriving its negation \( \neg P \). This, as it turns out, is easily done by combining \( P \) with the starting conditional \( P \rightarrow \neg P \), so on line 4 we derive \( \neg P \) and complete the contradiction and hence the indirect proof, allowing us to cancel the Show of line 2 and end the proof.

The inference from \( P \rightarrow \neg P \) to \( \neg P \) captures a pattern we see in self-refuting sentences. Consider, for example, the claim that

Every sentence that starts with the letter 'E' is false.

Suppose we want to show that this claim is false. One way we can do so is to point out that since that very sentence starts with the letter 'E', if it is a correct claim, it implies that it itself is false -- that is \( P \rightarrow \neg P \), where \( P \) is the claim in question. From this, we can conclude that the claim is in fact false, since the hypothesis that it is true is immediately self-refuting. Notice that this way of showing the claim false has a certain logical directness and unassailability that we don't get just by bringing up some specific example of a true sentence starting with the letter 'E' (Egglpasts are purple, perhaps). Philosophers sometimes try to derive quite powerful conclusions from this kind of self-refutation argument. It has been suggested, for example, that the claim that Truth is relative is self-refuting, since if it were true, its own truth would have to be relative. Much later, we'll see a variant on this self-refutation reasoning play a central role in Gödel's First Incompleteness Theorem, which is one of the most important results established by modern logic.

Now let's look at a slightly less trivial example. We will prove:

- \( P \rightarrow (R \& S) \), \( \neg P \rightarrow T \), \( (R \& P) \rightarrow (S \& \neg T) \), ∴ T

The proof proceeds as follows:

| (1) \( P \rightarrow (R \& S) \) | A |
| (2) \( \neg P \rightarrow T \) | A |
| (3) \( (R \& P) \rightarrow (\neg S \& \neg T) \) | A |
| (4) Show \( T \) | |
| (5) Show \( \neg P \) | |
| (6) \( P \) | AIP |
| (7) R \& S | \( \rightarrow E, 1, 6 \) |
| (8) R | &E, 7 |
| (9) R \& P | &I, 5, 6 |
| (10) \( \neg S \& \neg T \) | \( \rightarrow E, 3, 9 \) |
| (11) \( \neg S \) | &E, 10 |
| (12) S | &E, 7 |
| (13) T | \( \rightarrow E, 2, 5 \) |

This proof starts with a direct proof (triggered by the Show of \( T \) on line 4), but then immediately launches into an indirect subproof to conclude \( \neg P \). Once both \( \neg S \) and S are derived on lines 11 and 12, the indirect proof concludes, and the \( \neg P \) derived from it can be combined with the conditional \( \neg P \rightarrow T \) to get the final goal of \( T \).

**Question**: How do you know to start an indirect proof in this example?

**Answer**: The primary goal of the proof is to derive \( T \), so we look at the information we have in the three assumptions and see if there is any obvious route to reaching that goal. The conditional \( \neg P \rightarrow T \) is an obvious route -- if we knew that \( \neg P \),
we could use \( \rightarrow \)E to derive \( \neg \)P. Thus we set a secondary goal of deriving \( \neg \)P. Since \( \neg \)P is a negated sentence, indirect proof is an obvious way of deriving it. Thus we introduce the Show line on line 5, triggering an indirect proof of \( \neg \)P.

**Question:** How do you know what contradiction to derive to complete the subproof? Why S and \( \neg \)S in this case, rather than R and \( \neg \)R, or P and \( \neg \)P, or even \( \neg \)R \& S and \( \neg \)(R \& S)?

**Answer:** This is a good question, and unfortunately there's no simple answer to it. Knowing what contradiction to try for is often one of the most difficult aspects of succeeding in an indirect proof. Any contradiction at all will suffice, so there are an infinite number of choices available. About the best that can be done is to point out that:

- It's useful if half of the contradiction is already available, so a good place to start is by looking at what you already know and seeing if any of it can perhaps serve as half of a contradiction. In the proof above, I can see by looking over my assumptions that both S and T are available in negated and unnegated form in the consequents of conditionals, so both of those are good bets for places to look for a contradiction.
- In most cases, the extra assumption (taken by AIP) will be crucial in obtaining at least one wing of the contradiction, so it's often helpful to look and see what inferential power you gain through that assumption. In the proof above, the AIP of P puts me in a position to derive R \& S from the first conditional and gets me on the way to being able to derive \( \neg \)S \& \( \neg \)T from the third conditional, so it's not a bad bet that S will be involved. (It's also worth noting here that I already plan to use the second conditional -- \( \rightarrow \)P \( \rightarrow \)T -- to derive T in the end, so there's a good chance that it will not be involved in deriving the contradiction. While it's not an infallible rule, it's frequently the case that each assumption has one particular use in the proof.)

So, we now have three proof strategies available:

- **Direct Proof:** In direct proof, a Show line declares an intention to show some claim. That claim can be anything the prover desires. In order to complete the triggered direct subproof, some later line of the proof must contain the claim the Show line declares an intention to show.

- **Conditional Proof:** In conditional proof, a Show line declares an intention to show a claim whose main connective is a conditional. The triggered conditional subproof then begins by adding as a new assumption (using ACP) the antecedent of the conditional to be proved. The subproof is then completed (and the Show line cancelled) when the consequent of the conditional has been derived.

- **Indirect Proof:** In indirect proof, a Show line declares an intention to show a claim whose main connective is a negation. The triggered indirect subproof then begins by adding as a new assumption (using AIP) the unnegated form of the sentence to be shown. The subproof is then completed (and the Show line cancelled) when any contradiction (i.e., two lines, one of which is the negation of the other) has been derived.

Remember that it is very important that *every time you write a Show line in a proof, you must select some proof strategy for the subproof triggered by that Show line.* It never makes sense simply to declare an intention to Show -- you must have a type of proof selected. Writing a Show line without having a proof strategy in mind is a frequent source of trouble for students trying to learn the proof system.

It's also crucial that the proof strategy you select be one which is applicable to the Show line you've selected. The constraints are very simple:

- Direct proof can be used to prove any kind of sentence.
- Conditional proof can only be used to prove a sentence whose main conditional is a conditional.
- Indirect proof can only be used to prove a sentence whose main conditional is a negation.

As a quick test, try determining what kind of proof strategy can be be invoking for each of the following Show lines, and then compare your answers to the correct answers:

- Show \( \neg \)(P \( \rightarrow \)Q)
- Show \( \neg \)(P \( \rightarrow \)(Q \( \rightarrow \)R))
- Show \( \neg \)P \( \leftrightarrow \)Q
- Show (P \( \rightarrow \)Q) \( \leftrightarrow \)(R \& S)
- Show \( \neg \)(P \& Q)
- Show \( \neg \)(P \& \( \neg \)Q)
- Show ((P \( \rightarrow \)Q) \( \vee \)R) \( \rightarrow \)(Q \( \leftrightarrow \)(R \( \rightarrow \)S))

If you got any of these wrong, I very strongly recommend you go back over the constraints on invoking proof strategies, and make sure you've got everything clear.
Now that we've got the basics of indirect proof down, let's look at some more sophisticated uses of indirect proof, by considering the way in which it can be combined with conditional proof. We'll first prove the following result:

- \( P \rightarrow R \), \( Q \rightarrow \neg R \), \( \therefore P \rightarrow \neg Q \)

(which can be seen as saying if contradictory results follow from two different hypotheses, then if one hypothesis is true, the other is false.) The proof proceeds as follows:

1. \( P \rightarrow R \) A
2. \( Q \rightarrow \neg R \) A
3. \( \text{Show } P \rightarrow \neg Q \)
4. \( P \) ACP
5. \( \text{Show } \neg Q \)
   - \( Q \) AIP
   - \( R \rightarrow \neg R \rightarrow E, 1,4 \)
   - \( \neg R \rightarrow \neg R \rightarrow E, 2,6 \)

Here we start with a conditional proof, which gives us the auxiliary hypothesis \( P \). Our goal then shifts to proving the consequent \( \neg Q \), so we begin an indirect proof and take a second auxiliary hypothesis \( Q \). These two auxiliary hypotheses, combined with the two starting conditionals, allow us to derive a contradiction, which completes the indirect proof.

Completing the indirect proof then cancels the \text{Show} of line 5, and hence derives the consequent of the conditional of line 3, and thus also completes the conditional proof.

So we can have an indirect proof nested within a conditional proof. We can also do it the other way around, and have a conditional proof nested within an indirect proof, as in the proof of:

- \( P \rightarrow (P \rightarrow P) \), \( \therefore \neg P \)

which proceeds as follows:

1. \( P \rightarrow (P \rightarrow P) \) A
2. \( \text{Show } \neg P \)
3. \( P \) AIP
4. \( (P \rightarrow P) \rightarrow E, 1,3 \)
5. \( \text{Show } P \rightarrow P \)
   - \( P \) ACP

This proof can be confusingly abrupt in its ending, so it's worth taking the time to make sure it's clear why it has been successfully completed. The main strategy of the proof is the indirect proof, triggered by the \text{Show} on line 2. So as soon as a contradiction is derived, we are done. In line 4, we derive \( (P \rightarrow P) \), which will serve as one half of our contradiction. To complete the contradiction, we need \( P \rightarrow P \). Thus we begin a conditional proof to obtain this result. Of course (as we saw earlier), deriving \( P \rightarrow P \) is an extremely trivial task -- the ACP on line 6 serves as both antecedent and consequent, and brings that subproof to a quick close. But the conclusion of that subproof establishes the other half of the contradiction, and hence serves to complete the indirect proof as well (and hence the proof as a whole).

The Minus-First Rule: A Slight Side Trip [Next]

We're now going to show that, for boring technical reasons, we need to add one more proof rule to the system. This rule isn't really a substantive one, it's just one required for bookkeeping reasons. To see why we need it, suppose we want to prove the following result:

- \( P \rightarrow Q \), \( \therefore \neg Q \rightarrow \neg P \)

This is a very important result, and one we'll appeal to many times in constructing proofs -- it's the result of the contrapositive, which allows us to use conditionals in an inferentially "backward" manner, as well as forward.

This proof runs quite straightforwardly most of the way through. We get:

1. \( P \rightarrow Q \) A
2. \( \text{Show } \neg Q \rightarrow \neg P \)
3. \( \neg Q \) ACP
4. \( \text{Show } \neg P \)
   - \( P \) AIP
   - \( Q \rightarrow \neg E, 1,5 \)
We're trying to show \( \neg Q \rightarrow \neg P \), so we start with a conditional proof. The conditional proof gives us \( \neg Q \), and requires us to show \( \neg P \). Since \( \neg P \) is negated, we use indirect proof to show it. The indirect proof gives us \( P \) to work with, and obliges us to derive a contradiction. \( P \), combined with our initial \( P \rightarrow Q \), gives us \( Q \), which looks like a promising half of a contradiction. All we need now is \( \neg Q \). Even better, we've already got \( \neg Q \), back on line 3. The only problem is that \( \neg Q \) isn't inside the indirect proof, so the indirect proof doesn't contain a contradiction, and thus doesn't meet the completion conditions.

What we need, then, is a rule which will allow us to move the \( \neg Q \) down into the indirect proof. This ought to be permissible, since downward (but not upward or sideways) transmission of information is fine. We thus add a rule of Reiteration, or \( R \), which allows us simply to repeat a line that we had earlier derived, so long as that line is either in the same subproof or in a higher subproof (i.e., so long as the reiteration does not involve upward or sideways information transmission). The rule of Reiteration has the following stunningly boring input-output structure:

\[
\begin{array}{c}
\text{R:} \\
\Phi \\
\end{array}
\]

The input is exactly the same as the output, because all we are doing is reiterating.

With the \( R \) rule added, we can successfully complete our proof above:

1. \( P \rightarrow Q \) A
2. Show \( \neg Q \rightarrow \neg P \)
3. \( \neg Q \) ACP
4. Show \( \neg P \)
   - 5. \( P \) AIP
   - 6. \( Q \) →E, 1,5
   - 7. \( \neg Q \) R, 3

The last step in the proof, then, is to reiterate the \( \neg Q \) we had obtained earlier, and thus complete the contradiction and finish the proof.

We could have gotten by without a reiteration rule, at the cost of some decrease in readability. We could have allowed indirect proofs to count as finished as soon as there was a contradiction anywhere either in the indirect proof itself or in any higher-level proof (and similarly allowed a conditional proof to count as finished as soon as the consequent was anywhere either in the conditional proof itself or in any higher-level proof). The technical result would have been identical, but it would have at times been extremely mysterious why a subproof ended when it did.

**Back to Indirect Proofs**

OK, now back to the indirect proofs. We'll look at a few more examples, these a step or two more elaborate, and then go over some questions about the strategy of indirect proof.

First, let's look at an example that has multiple levels of nesting of indirect and conditional proofs. We'll prove the following result:

\[ R \rightarrow ((U \rightarrow (P \lor Q)) \rightarrow S), T \rightarrow \neg(R \rightarrow S). \therefore (P \lor Q) \rightarrow \neg T \]

The proof proceeds as follows:

1. \( R \rightarrow ((U \rightarrow (P \lor Q)) \rightarrow S) \) A
2. \( T \rightarrow \neg(R \rightarrow S) \) A
3. Show \( (P \lor Q) \rightarrow \neg T \)
4. \( P \lor Q \) ACP
   - 5. Show \( \neg T \)
     - 6. \( T \) AIP
     - 7. \( \neg(R \rightarrow S) \) →E, 2,6
     - 8. Show \( R \rightarrow S \)
       - 9. \( R \) ACP
         - 10. Show \( U \rightarrow (P \lor Q) \)
           - (11) \( U \) ACP
           - (12) \( P \lor Q \) R, 4
             - 13. \( (U \rightarrow (P \lor Q)) \rightarrow S \) →E, 1,9
Here we start with a conditional proof (to establish \((P \lor Q) \rightarrow \neg T\)), and then move to an indirect proof (to establish \(\neg T\)), and then to a further conditional proof (to get \(R \rightarrow S\), as part of the contradiction that will complete the indirect proof), and finally during that last conditional proof do one more conditional proof to get \(U \rightarrow (P \lor Q)\), which is crucial for getting the consequent \(S\) which completes the penultimate conditional proof.

Next, we'll look at a proof which contains multiple levels of indirect proof. We'll prove the following claim:

- \(P \rightarrow Q\)
- \(\neg R \rightarrow \neg Q\)
- \(R \rightarrow S\)
- \(T \rightarrow \neg S\)
- \(\neg T \rightarrow \neg P\)

\(\therefore \neg P\)

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise/Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(P \rightarrow Q) A</td>
</tr>
<tr>
<td>2</td>
<td>(\neg R \rightarrow \neg Q)</td>
</tr>
<tr>
<td>3</td>
<td>(R \rightarrow S)</td>
</tr>
<tr>
<td>4</td>
<td>(T \rightarrow \neg S)</td>
</tr>
<tr>
<td>5</td>
<td>(\neg T \rightarrow \neg P) A</td>
</tr>
<tr>
<td>6</td>
<td>Show (\neg P)</td>
</tr>
<tr>
<td>7</td>
<td>(P) AIP</td>
</tr>
<tr>
<td>8</td>
<td>(Q) \rightarrow E, 1,7</td>
</tr>
<tr>
<td>9</td>
<td>Show (\neg R)</td>
</tr>
<tr>
<td>10</td>
<td>(R) AIP</td>
</tr>
<tr>
<td>11</td>
<td>(S) \rightarrow E, 3,10</td>
</tr>
<tr>
<td>12</td>
<td>Show (\neg T)</td>
</tr>
<tr>
<td>13</td>
<td>(T) AIP</td>
</tr>
<tr>
<td>14</td>
<td>(\neg S) \rightarrow E, 4,13</td>
</tr>
<tr>
<td>15</td>
<td>(S) R, 11</td>
</tr>
<tr>
<td>16</td>
<td>(\neg P) \rightarrow E, 5,12</td>
</tr>
<tr>
<td>17</td>
<td>(P) R, 7</td>
</tr>
<tr>
<td>18</td>
<td>(\neg Q) \rightarrow E, 2,9</td>
</tr>
</tbody>
</table>

This proof contains three nested indirect proofs. Although the logic of the proof may appear somewhat complex, really each indirect proof is fairly immediately triggered by thinking about the goals of the proof. The first subproof, obviously, is triggered just by the stated goal of the proof. We're now out to find a contradiction. We get \(Q\) right off, so \(\neg Q\) is an obvious place to look for the contradiction. From \(\neg R \rightarrow \neg Q\), we see that if we had \(\neg R\), we could get \(\neg Q\) -- this motivates the second subproof, trying to get that \(\neg R\). That subproof quickly gives us \(S\), so now we need \(\neg S\) to get what we want. At this point we see that if we had \(\neg T\), we could get \(\neg S\), and this triggers the third subproof, which leads directly to a contradiction.

(Note that we could have tried doing an indirect proof directly to get \(\neg Q\), taking \(Q\) for AIP. However, this approach wouldn't have led anywhere, because none of our premises combine with \(Q\) to yield useful results. Similarly, we could have tried to get \(\neg S\) directly through indirect proof, taking \(S\) for AIP. But this also would have failed, because \(S\) would not combine with any of our assumptions. Sometimes it takes a few tries to find the right path through the proof.)

Another Kind of Use of Indirect Proof, and the Relation Between It and the \(\neg \neg\) Rule [Next]

Earlier we proved what I called the self-refutation result that from \(P \rightarrow \neg P\) one can derive \(\neg P\). Suppose we want to prove the corresponding self-affirmation result -- that from \(\neg P \rightarrow P\), one can conclude \(P\). This reasoning is used in the following argument:

Consider the claim There is at least one false claim. If this claim is false, then it must be correct, because it will itself provide the one false claim. Thus it must be correct.

How can we produce a formal proof of this claim? It looks like we want to do some sort of indirect proof -- the thought is that if \(P\) is false, it is also true, which is a contradiction, forcing us to reject the assumption that \(P\) is false and adopt instead the assumption that \(P\) is true. But things aren't set up right for an indirect proof here -- consider how the proof would start:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise/Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\neg P \rightarrow P) A</td>
</tr>
<tr>
<td>2</td>
<td>Show (P)</td>
</tr>
</tbody>
</table>
We can't invoke indirect proof in response to the Show of line 2, because the content of the Show line does not have negation as its main connective.

However, we can make things work out if we combine indirect proof with the \( \neg \neg \) rule. Instead of trying to show \( P \), which blocks us from using indirect proof, we'll try instead to show \( \neg \neg P \), which is eligible for indirect proof (since it has a negation as its main connective). Once we succeed in showing \( \neg \neg P \), we'll use \( \neg \neg \) to transform it into \( P \). The proof thus proceeds as follows:

\[
\begin{align*}
(1) & \quad \neg P \rightarrow P \quad A \\
(2) & \quad \text{Show } P \\
(3) & \quad \text{Show } \neg \neg P \\
(4) & \quad \neg P \quad \text{AIP} \\
(5) & \quad P \quad \text{E, 1, 4} \\
(6) & \quad P \quad \neg \neg 3 \\
\end{align*}
\]

This technique of combining indirect proof with \( \neg \neg \) effectively allows us to use indirect proof to establish unnegated claims as well as negated claims, by using indirect proof to get the double-negated form of what we want and then using \( \neg \neg \) to remove that double negation. This technique often comes in useful. Here's another example -- we'll prove:

- \( (P \rightarrow Q) \rightarrow P \quad \therefore \quad P \)

The proof proceeds as follows:

\[
\begin{align*}
(1) & \quad (P \rightarrow Q) \rightarrow P \quad A \\
(2) & \quad \text{Show } P \\
(3) & \quad \text{Show } \neg \neg P \\
(4) & \quad \neg P \quad \text{AIP} \\
(5) & \quad \text{Show } \neg (P \rightarrow Q) \\
(6) & \quad P \rightarrow Q \quad \text{AIP} \\
(7) & \quad P \quad \text{E, 1, 6} \\
(8) & \quad \neg P \quad \text{R, 4} \\
(9) & \quad \text{Show } P \rightarrow Q \\
(10) & \quad P \quad \text{ACP} \\
(11) & \quad \text{Show } \neg \neg Q \\
(12) & \quad \neg Q \quad \text{AIP} \\
(13) & \quad P \quad \text{R, 10} \\
(14) & \quad \neg P \quad \text{R, 8} \\
(15) & \quad Q \quad \neg \neg 11 \\
(16) & \quad P \quad \neg \neg 3 \\
\end{align*}
\]

This proof is quite a complex one, and has a number of less-than-obvious moves in it. Of most immediate interest, it twice uses the technique of proving an unnegated claim by using indirect proof to establish the doubly-negated claim and then using \( \neg \neg \). This technique is used on line 3 to establish \( \neg \neg P \), and again on line 11 to establish \( \neg \neg Q \).

This proof also makes the somewhat odd move of performing an indirect proof (from lines 12 to 14) in which both halves of the contradiction come from outside the indirect proof, and in which the assumption for indirect proof is never in fact used. The subproof runs this way because we have already established \( \neg P \), and the falsity of \( P \) is enough to establish the truth of the conditional \( P \rightarrow Q \). In fact, this is a useful result that gets applied not infrequently in proofs:

- \( \neg P \quad \therefore \quad P \rightarrow Q \)

\[
\begin{align*}
(1) & \quad \neg P \\
(2) & \quad \text{Show } P \rightarrow Q \\
(3) & \quad P \quad \text{ACP} \\
(4) & \quad \text{Show } \neg \neg Q \\
(5) & \quad \neg Q \quad \text{AIP} \\
(6) & \quad P \quad \text{R, 3} \\
(7) & \quad \neg P \quad \text{R, 1} \\
(8) & \quad Q \quad \neg \neg 4 \\
\end{align*}
\]
As you’ve probably noticed, it’s a bit of a nuisance to have to keep introducing and then eliminating double negatives any
time you want to use indirect proof to establish an unnegated claim. Since we’ve now seen that any time we want to use
indirect proof to establish an unnegated claim \( \Phi \) we can do so using the following proof scheme:

\[
\begin{array}{c}
(1) \text{Show } \Phi \\
(2) \text{Show } \neg \neg \Phi \\
(3) \neg \Phi \\
(4) \Theta \\
(5) \neg \Theta \\
(6) \Phi \\
\end{array}
\]

(where the ellipsis contains the process of deriving the contradiction to complete the indirect proof), we will from now on
allow ourselves to drop the steps in lines 2 and 6, and directly use indirect proof to establish unnegated claims. We thus
introduce the following alternative form of indirect proof:

**Indirect Proof (Alternative Form):** In the alternative form of indirect proof, a Show line declares an intention to
show any claim. The triggered indirect subproof then begins by adding as a
new assumption (using AIP) either (a) the unnegated form of the sentence to
be shown, if that sentence has a negation as its main connective, or (b) the
negation of the sentence to be shown, if that sentence does not have a
negation as its main connective. The subproof is then completed (and the
Show line cancelled) when any contradiction (i.e., two lines, one of which is
the negation of the other) has been derived.

We won’t make any distinction between the standard and alternative forms of indirect proof when we do proofs; we’ll
simply give a Show line and then introduce the appropriate auxiliary assumption using AIP.

With the alternative form of indirect proof, the proof of:

- \((P \rightarrow Q) \rightarrow P, \therefore P\)

simplifies into:

\[
\begin{array}{c}
(1) (P \rightarrow Q) \\
(2) \text{Show } P \\
(3) \neg P \\
(4) \text{Show } \neg (P \rightarrow Q) \\
(5) P \rightarrow Q \\
(6) P \\
(7) \rightarrow E, 1,5 \\
(8) \neg P \\
(9) \text{Show } P \rightarrow Q \\
(10) P \\
(11) \text{Show } Q \\
(12) \neg Q \\
(13) P \\
(14) \neg P \\
\end{array}
\]

Once we add the alternative version of indirect proof, and thus avoid the need of routing obliquely through the use of \( \neg \neg \)
it’s natural to ask what role the rule of \( \neg \neg \) will play in our proof system. Recall that \( \neg \neg \) actually licenses two types of
inference – one from \( \Phi \) to \( \neg \neg \Phi \), and one from \( \neg \neg \Phi \) to \( \Phi \). As it turns out, the first of these two inferences isn’t necessary
even before we add the alternative version of indirect proof. Using standard indirect proof, we can always add a double
negative to a sentence. Thus consider the following proof of:

- \( P, \therefore \neg \neg P \)

\[
\begin{array}{c}
(1) P \\
(2) \text{Show } \neg \neg P \\
(3) \neg P \\
(4) P \\
\end{array}
\]

However, the same technique won’t work to capture the inference from \( \neg \neg P \) to \( P \) – that is, the elimination of the double
negative. If we try to start a proof the same way:

\[
\begin{array}{c}
(1) \neg \neg P \\
(2) \text{Show } P \\
(3) \neg P \\
(4) P \\
\end{array}
\]
we immediately come up against the problem that we can't use indirect proof to show $P$, since it is unnegated. However, with the alternative form of indirect proof, we can complete this proof quite easily:

| (1)               | A               |
| (2) Show $P$          | A               |
| (3)                | A               |
| (4)               | A               |

Since $\neg P$ and $\neg \neg P$ are a contradictory pair, this completes the indirect proof.

What we have, then, is really a total of four rules:
- $\neg\neg$ in the introductory direction, from $\Phi$ to $\neg\neg\Phi$
- $\neg\neg$ in the eliminatory direction, from $\neg\neg\Phi$ to $\Phi$
- Standard indirect proof, for concluding $\neg\Phi$
- Alternative indirect proof, for concluding $\Phi$

It would suffice to have either (a) standard indirect proof, plus $\neg\neg$ in the eliminatory direction, or (b) both standard and alternative indirect proof. Just to make things easy on ourselves, however, we will allow ourselves all four of these techniques.

**Questions About Indirect Proof** [Next]

Let's now go over some common questions about indirect proof.

**Question #1**: If I derive $P \rightarrow Q$ and $P \rightarrow \neg Q$ within an indirect proof, does that count as deriving a contradiction (and hence finishing the indirect proof)?

**Answer**: No. $P \rightarrow Q$ and $P \rightarrow \neg Q$ are not contradictories. Both of them can be true if $P$ is false. The only thing that counts as deriving a contradiction is getting a sentence and the negation of that sentence.

**Question #2**: OK, what if I derive $Q \land R$ and $T \land \neg R$? Does that count as a contradiction? Those can't both be true.

**Answer**: No, it still doesn't count as a contradiction. It's true that those two sentences can't both be true, and a more liberalized version of indirect proof could allow that pair to complete the subproof. Our system, however, follows the strict requirement that you must get a sentence and its negation, which you don't have in this situation. However, once you've got $Q \land R$ and $T \land \neg R$, you can easily get a strict contradiction, by applying $\&E$ to extract $R$ and $\neg R$.

**Question #3**: So, am I supposed to be using the standard form of indirect proof, or the alternative form?

**Answer**: Use the alternative form. Notice that everything that can be done with the standard form can also be done with the alternative form -- the standard form is just a special case of the alternative form. I've made the distinction between the two just because the book we're working with in class initially sets up indirect proof in what I've called the standard form. If I weren't working with that book, I'd just go straight to the alternative form and not mention the standard form at all.

**Question #4**: Sometimes the contradiction in an indirect proof seems to have nothing to do with the assumption for indirect proof, like in the following:

| (1) $P$               | A               |
| (2) $\neg P$          | A               |
| (3) Show $Q$          | A               |
| (4) $\neg Q$          | A               |
| (5) $P$               | R               |
| (6) $\neg P$          | R               |

Here the contradiction comes completely from outside the indirect proof, and is just imported using the $R$ rule. Is this legitimate?
Answer: Yes. So long as there's some contradiction within the indirect proof, it doesn't matter where it comes from. In most cases, at least half of the contradiction will come from the AIP (since it is assuming a true claim to be false that leads to the contradiction), but in this case we already had a contradiction in our starting assumptions, so all we had to do was import it to the indirect proof.

Question #5: But if the previous is a legitimate proof, so is the following:

\[
\begin{align*}
(1) & \quad P & A \\
(2) & \quad \neg P & A \\
(3) & \quad \text{Show } \neg Q & \\
(4) & \quad Q & \text{AIP} \\
(5) & \quad P & R, 1 \\
(6) & \quad \neg P & R, 2 \\
\end{align*}
\]

even though this proof reaches the exact opposite conclusion from the very same assumptions. In fact, it looks like we can derive any conclusion at all from these assumptions, since the conclusion (and hence the AIP which is the opposite of the conclusion) plays no role in the proof. Can that be right?

Answer: Yes, this is right. If you've got contradictory premises, you can derive any conclusion at all. This is once again the principle ex falso quodlibet, which we saw earlier from the semantic side (where it was justified by showing that if our premises included a contradiction, then there was no interpretation making all the premises true, and hence automatically no interpretation making all the premises true and the conclusion false). Now we've seen that the same principle can be established from the syntactic side using the proof system. As before, although it looks odd, it's a safe principle because you can never really have true premises which are contradictory, so you don't need to worry about what follows from contradictory premises. As we'll see much later, it's possible to adopt an alternative view of logic on which ex false quodlibet is not a correct principle, so if you're really bothered by this feature of classical logic, you'll eventually learn a way around it (although there's a not inconsiderable price to be paid for avoiding it).

The Seventh Rule: Disjunction Introduction [Next]

Now that we've finally finished dealing with negation, we're ready to move on to a new connective: the disjunction. You'll notice that I've departed from my previous practice of giving the elimination rule before the introduction rule. That's because in the case of disjunction, the introduction rule is quite a bit simpler than the elimination rule, so I thought it would be easier to tackle it first.

A rule of disjunction introduction, or $\lor I$, should enable us to reach conclusions of the form $\Phi \lor \Theta$, so we need to think about what sort of input would be needed to allow us to reach our desired conclusion. Since the disjunction is true if either disjunct is true, we can easily see that having either disjunct as input would suffice for reaching our desired conclusion. Thus we have our official input-output formulation of $\lor I$:

\[
\begin{align*}
\Phi & : \quad \text{or} \\
\Theta & : \\
\Phi \lor \Theta & : \quad \Phi \lor \Theta \\
\Theta \lor \Phi & : \quad \Theta \lor \Phi
\end{align*}
\]

Recall that I talked earlier (when giving the conjunction rules) about the downhill flow of information, using this metaphor to explain why it's easier to formulate a rule of conjunction elimination than a rule of conjunction introduction. The same metaphor applies here, but in the opposite direction. A rule of disjunction introduction is easy, because we are allowing information to flow from the higher position of $\Phi$ (in which we know that one particular claim is true) to the lower position of $\Phi \lor \Theta$ (in which we know only that one of two claims is true, without knowing which). And, as we'll see in a bit, a rule of disjunction elimination (which, prima facie, would allow for uphill information flow) is quite difficult to formulate.

Let's look at a few examples of $\lor I$. I won't go through a lot of examples, because the rule is not that complicated. First, we'll prove the following useful result:

- $(P \lor Q) \rightarrow R, \therefore P \rightarrow R$

The proof proceeds as follows:

\[
\begin{align*}
(1) & \quad (P \lor Q) \rightarrow R & A \\
(2) & \quad \text{Show } P \rightarrow R & \\
(3) & \quad P & \text{ACP} \\
(4) & \quad P \lor Q & \lor I, 3 \\
(5) & \quad R & \rightarrow E, 1,4
\end{align*}
\]
Here, $\lor$ is used on line 4 to move from $P$ to $P \lor Q$, in order to let us use the conditional we have as our starting assumption.

Next, we'll prove the following result:

$P \rightarrow Q, (R \lor Q) \rightarrow S, S \rightarrow \neg P, \therefore \neg P$

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise/Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$P \rightarrow Q$</td>
</tr>
<tr>
<td>2</td>
<td>$(R \lor Q) \rightarrow S$</td>
</tr>
<tr>
<td>3</td>
<td>$S \rightarrow \neg P$</td>
</tr>
<tr>
<td>4</td>
<td>Show $\neg P$</td>
</tr>
<tr>
<td>5</td>
<td>$P$</td>
</tr>
<tr>
<td>6</td>
<td>$Q \rightarrow \neg E, 1, 5$</td>
</tr>
<tr>
<td>7</td>
<td>$(R \lor Q) \lor I, 6$</td>
</tr>
<tr>
<td>8</td>
<td>$S \rightarrow \neg E, 2, 7$</td>
</tr>
<tr>
<td>9</td>
<td>$\neg P \rightarrow \neg E, 3, 8$</td>
</tr>
</tbody>
</table>

The rule of $\lor$I, used here on line 7, is serving basically the same function as it did in the previous proof -- weakening a piece of information to allow it to serve as the antecedent of a conditional in an application of $\rightarrow$E. Note, however, that here we add the new disjunct to the left side, rather than the right side, of the disjunction.

Finally, let's look at a rather different example, in which we prove the extremely useful result:

$\therefore P \lor \neg P$

This result is an instance of the law of the excluded middle, which states that, for any sentence $\Phi$, either $\Phi$ or $\neg \Phi$ is true.

As we'll see below, being able to derive an instance of the law of the excluded middle is occasionally a quite useful proof technique. The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise/Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Show $P \lor \neg P$</td>
</tr>
<tr>
<td>2</td>
<td>$\neg (P \lor \neg P)$ AIP</td>
</tr>
<tr>
<td>3</td>
<td>Show $\neg P$</td>
</tr>
<tr>
<td>4</td>
<td>$P$ AIP</td>
</tr>
<tr>
<td>5</td>
<td>$P \lor \neg P \lor I, 4$</td>
</tr>
<tr>
<td>6</td>
<td>$\neg (P \lor \neg P) \rightarrow \neg E, 2$</td>
</tr>
</tbody>
</table>

Although relatively short, this proof is far from obvious in its structure. It's much less predictable in the way it proceeds than most of the proofs we've looked at so far, and (unlike many of the earlier proofs) is probably not something you'd come up with on your first try at it. We start off trying to prove the disjunction $P \lor \neg P$. Our new rule of $\lor$I would seem an obvious thing to use here, but to use it we'd need to have either $P$ or $\neg P$, which we don't have. (The point, after all, is to show that the disjunction $P \lor \neg P$ is true regardless of whether $P$ is true or false.) So we have to find a different approach.

We thus use the following very useful bit of advice for constructing proofs:

• When in doubt, try an indirect proof.

Now that we've admitted the alternative form of indirect proof, indirect proof can be invoked for any Show line, and it has the advantage of giving you another assumption to work with.

Using indirect proof, we then add the assumption $\neg (P \lor \neg P)$. Now we need to derive a contradiction. Perhaps surprisingly, it turns out that the best way to get that contradiction is by deriving $P \lor \neg P$ to contradict $\neg (P \lor \neg P)$. This is surprising because, of course, $P \lor \neg P$ was what we wanted in the first place, so it seems a bit odd to be deriving it within the indirect proof -- why not derive it directly? The answer is that the auxiliary premise $\neg (P \lor \neg P)$ will give us the extra inferential power to derive $P \lor \neg P$.

Obviously it would be pointless to start another indirect proof to get $P \lor \neg P$, so this time we are going to use $\lor$I. To do so, we need either $P$ or $\neg P$. In the proof given above, we use indirect proof to get $\neg P$ -- this proceeds by taking $P$ for $\lor$I, and then using $\lor$I to make that $P$ into $P \lor \neg P$, and then contradicting that with the $\neg (P \lor \neg P)$ we already had. (Notice that we cannot use the $P \lor \neg P$ that we derive on line 5 to contradict $\neg (P \lor \neg P)$ on line 2 and complete the first indirect proof, because line 5 is sealed in the subproof by the ban on upward transmission of information.) We could also have derived $P$ in the same manner.

Once we've derived $\neg P$, we use $\lor$I again to change it into $P \lor \neg P$, which contradicts $\neg (P \lor \neg P)$, and thus completes the first indirect proof. Notice that $P \lor \neg P$ occurs three times in the proof -- on lines 1, 5, and 7. It is only the occurrence on
line 1, after the Show has been cancelled, that completes the proof. The other two are sealed in subproofs. It's not an uncommon pattern in indirect proofs that the claim you want to prove shows up multiple times before it really counts as proved. It's a very good exercise to make sure you are perfectly clear on why the occurrences of $P \lor \neg P$ on lines 5 and 7 do not count as completing the proof.

Let's go over a few common questions about $\lor I$ before we move on to disjunction elimination.

Question #1: Why do you cite only one input line when you use $\lor I$? If you're deriving $P \lor Q$, don't you need an input line for $P$ and an input line for $Q$?

Answer: No. Since a disjunction requires the truth of only one disjunct for its truth, we don't need to have established both $P$ and $Q$ earlier in the proof. If we've already established $P$, then we can safely conclude $P \lor Q$, whether we know $Q$ or not (indeed, whether $Q$ is true or not). Similarly if we've already established $Q$. $\lor I$ requires only a single input.

Question #2: Sometimes $\lor I$ seems to bring stuff in out of nowhere, like in the following:

| (1) $P$ | $\lor I$, 1 |
| (2) $P \lor \neg (R \leftrightarrow (Q \leftrightarrow S \land (T \lor \neg U)))$ | $\lor I$, 1 |

None of this stuff about $Q$ and $R$ and $S$ and $T$ and $U$ was in the proof before -- it just appears out of nowhere when $\lor I$ is used. Is that legitimate?

Answer: Yes. $\lor I$ can introduce anything you like as the second disjunct. So if we've got $P$ already in the proof, we can use $\lor I$ to obtain $P \lor \Phi$, or $\Phi \lor P$, for any sentence $\Phi$ we want. $\Phi$ can be as complex (or as simple) as we want, and it can have as little (or as much) bearing on what has gone before in the proof as we want. $\Phi$ is, in essence, junk information -- we're simply saying that since $P$ is true, either $P$ or this completely arbitrary other thing must be true.

Question #3: If I can introduce anything I want using $\lor I$, how do I know what to introduce? All the other rules told me exactly what output I got when I used them; I'm not sure what to do with the freedom that $\lor I$ gives me.

Answer: There's no simple answer to this, but the best guideline is to be thinking all the time about what you need to successfully complete the proof, and see if $\lor I$ fits into those needs. In the earlier derivation of $P \rightarrow R$ from $(P \lor Q) \rightarrow R$, for example, the initial conditional proof gives us $R$ as a success condition. The starting assumption of $(P \lor Q) \rightarrow R$ gives us a way of meeting that success condition, by using $\rightarrow E$ on the conditional. To do this, we need $P \lor Q$. Any time you need a disjunction, $\lor I$ is a good bet, so now we have a specific goal which can be met using $\lor I$, and it will be much easier to see what to do with the rule.

The Eighth Rule: Disjunction Elimination [Next]

We come now to the most complicated rule in the proof system -- the rule of disjunction elimination, or $\lor E$. Disjunction elimination is a difficult rule because it is trying to make a difficult inferential transition. A rule of $\lor E$ should allow us to take an input of the form $\Phi \lor \Theta$ and reach some non-disjunctive conclusion from it. But it's very hard to see how such an inferential move can conform to the principle of the downhill flow of information. A disjunction is already quite far downhill, since it tells us only that one of two things is true. It's hard to see how we can find an output location even farther downhill without moving to an even more disjunctive conclusion (like $\Phi \lor \Theta \lor \Psi$), which wouldn't really count as a disjunction elimination rule.

Let's start by looking at something that the rule of $\lor E$ is not (and should not be), in order to help bring out the difficulties. People just starting with the proof system are often tempted to have $\lor E$ work in what seems like the immediately obvious way:

$$
\lor E: \quad \Phi \lor \Theta \\
\Phi
$$

(Please note that this rule is in red because it is not correct.) Here $\lor E$ is acting directly as an elimination rule, simply disposing of one of the disjuncts. However, this "rule" leads to bad results. For example, using it we can prove the following "result":

- $P \lor Q, \therefore P$

with the following "proof":
But this result is a bad one, since an interpretation assigning $F$ to $P$ and $T$ to $Q$ will make the premise true and the conclusion false. The basic problem here is that you can't simply strike out one of the disjuncts, because for all you know that disjunct might be the true one.

So we need something more complex for our rule of $\lor E$. To help with working out what that something is, let's look at a sample bit of real-life reasoning involving disjunctions. Suppose I were to reason as follows:

Taxes will definitely go up after the 2000 presidential election. After all, either Gore or Bush will certainly win the election. If Gore wins the election, then he'll implement new social welfare programs, which will force the government to raise taxes to pay for them. On the other hand, if Bush wins the election, the stock market will crash and the tax revenue base will shrink, so the government will be forced to raise taxes just to keep income at its current level.

Notice how this argument proceeds. I start with a disjunctive assumption -- either Gore or Bush will win the 2000 presidential election. I end with a non-disjunctive conclusion derived from that assumption -- taxes will go up after the 2000 presidential election. I am able to reach this conclusion by what is sometimes called dilemma reasoning. The idea is that I forge an inferential path from each disjunct to the desired conclusion. First I show that Gore's being president leads to the raising of taxes. Then I show that Bush's being president leads to the raising of taxes. I still don't know which of these two will happen, of course, but I now know enough to conclude safely that taxes will be raised.

Pictorially, we might think of the situation like this:

This diagram represents paths the world could take. The first branching point represents the choice between Gore being president and Bush being president. The bottom point of the diagram represents taxes being raised. What the argument shows is that there is a required path from both of the branch points to this bottom point. If we now, as it were, start the world moving along this diagram, we can't predict exactly which path it will take -- but we can say with certainty that it will end up at the bottom point, since both available paths lead there. Thus we safely conclude that a raising of taxes will be a part of the world, and hence that taxes will be raised.

OK, now let's see how we can take this example and craft it into a rule. Our argument starts with some disjunctive assumption -- in the particular case above, that either Gore or Bush will win the election, and more generally, of the form $\Phi \lor \Theta$. It eventually reaches some conclusion which is not the same as either of the disjuncts -- in the particular case above, that taxes will be raised, and more generally, of the form $\Psi$. But the disjunction alone is not enough to reach this conclusion. We also need the inferential paths from each disjunct to the conclusion. Using conditionals as a method of displaying inferential paths, we add two more inputs -- one from each disjunct to the desired conclusion. In the particular case above, this means adding:

- If Gore wins the election, taxes will be raised.
- If Bush wins the election, taxes will be raised.

More generally, it means adding $\Phi \to \Psi$ and $\Theta \to \Psi$.

We are now ready to give the formal statement of the $\lor E$ rule. Its input-output format is:

- $\Phi \lor \Theta$
- $\Phi \to \Psi$  
- $\Theta \to \Psi$

$\lor E: \quad \Theta \to \Psi \quad \Phi \to \Psi \quad \Phi \lor \Theta \quad \Psi$
The $\lor E$ rule thus requires three separate inputs -- a disjunction, plus two conditionals expressing inferential connections between the disjuncts and the output. $\lor E$ is unique in being the only of our rules which mentions a connective other than the connective for which it is a rule. (We'll see later that it is possible to formulate a rule of $\lor E$ which makes no mention of any connectives other than disjunction -- although this rule is somewhat more theoretically elegant, it is also more awkward for beginners to use, so we sacrifice a bit of elegance for practicality here.)

Let's now look at a first example of this rule in action. We will prove the following claim:

$\varnothing \vdash P \lor Q, \neg P, \therefore Q$

The proof proceeds as follows:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Show $Q$</td>
</tr>
<tr>
<td>4</td>
<td>Show $P \rightarrow Q$</td>
</tr>
<tr>
<td>5</td>
<td>$P$ ACP</td>
</tr>
<tr>
<td>6</td>
<td>Show $Q$</td>
</tr>
<tr>
<td>7</td>
<td>$\neg Q$ AIP</td>
</tr>
<tr>
<td>8</td>
<td>$P$ R, 5</td>
</tr>
<tr>
<td>9</td>
<td>$\neg P$ R, 2</td>
</tr>
<tr>
<td>10</td>
<td>Show $Q \rightarrow Q$</td>
</tr>
<tr>
<td>11</td>
<td>$Q$ ACP</td>
</tr>
<tr>
<td>12</td>
<td>$Q \lor E, 1,4,10$</td>
</tr>
</tbody>
</table>

In this proof, we want to argue from the disjunction $P \lor Q$ to the specific conclusion $Q$ (with the aid of the second premise $\neg P$). We are thus required to beat an inferential path from each disjunct to the desired conclusion -- that is, to obtain the two conditionals $P \rightarrow Q$ and $Q \rightarrow Q$. The second of these, of course, is entirely trivially, and is obtained on lines 10 and 11. The first is somewhat more complicated, but is just another application of the technique which we saw earlier of deriving the truth of a conditional from the falsity of its antecedent. Once we've got both of these conditionals, we can combine them with the disjunction using $\lor E$ to get our final conclusion $Q$.

Note that the two conditionals we need are each obtained by invoking conditional proof. This is a frequent pattern in proofs involving $\lor E$ -- quite often an attempt to prove something from a disjunction will trigger the need to perform two conditional proofs, to get us the inputs we need for $\lor E$. Sometimes we will already have one or both of the conditionals to hand, but this is the exception rather than the rule.

Let's look at another example. This time we will prove an instance of the commutivity of disjunction, by proving the result:

$\varnothing \vdash P \lor Q, \therefore Q \lor P$

The proof proceeds as follows:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Show $Q \lor P$</td>
</tr>
<tr>
<td>4</td>
<td>Show $P \rightarrow (Q \lor P)$</td>
</tr>
<tr>
<td>5</td>
<td>$P$ ACP</td>
</tr>
<tr>
<td>6</td>
<td>$Q \lor P$ $\lor I, 4$</td>
</tr>
<tr>
<td>7</td>
<td>$Q$ ACP</td>
</tr>
<tr>
<td>8</td>
<td>$Q \lor P$ $\lor I, 7$</td>
</tr>
<tr>
<td>9</td>
<td>$Q \lor P$ $\lor E, 1,3,6$</td>
</tr>
</tbody>
</table>

Again we need to prove conditionals connecting each disjunct to our desired conclusion of $Q \lor P$. In this case, each conditional is established by using the rule of $\lor I$ to expand one of the two disjuncts into the entire disjunction. This proof follows a rather frequent pattern in proofs involving disjunctions:

- In a proof which has a disjunction as a premise and a disjunction as a conclusion, often you will use $\lor E$, and use one disjunct of the premise to prove one disjunct of the conclusion (and then use $\lor I$), and use the other disjunct of the premise to prove the other disjunct of the conclusion (and then use $\lor I$).

Next, let's prove a result that requires combining $\lor E$ with indirect proof. We'll prove the following result:

$\varnothing \vdash P \lor Q, \therefore \neg (P \land Q)$

(This result is an instance of the De Morgan equivalences). The proof proceeds as follows:
This proof has an indirect proof as its main strategy, but then uses the \( \lor \) \( \lor \) approach of dilemma reasoning to establish one half of the contradiction needed to conclude the indirect proof (namely, the \( \neg (P \land Q) \) half). Each branch of the dilemma reasoning then contains its own indirect proof.

Now let's go over some questions about \( \lor \lor \), and then come back and look at a few more complex examples.

**Question #1**: Are you sure I can't just use \( \lor \lor \) to infer \( P \) from \( P \) \( \lor \) \( Q \)?

**Answer**: Yes, quite sure. This is one of the most common mistakes in using \( \lor \lor \), and it's a very serious error -- a proof that misuses \( \lor \lor \) in this way will be a completely worthless proof (more to the point, it will likely get you very little credit on a problem set or test).

**Question #2**: When I do a proof using \( \lor \lor \), I keep getting the conclusion that I want over and over again. How can I know when I've gotten it enough times?

**Answer**: Yes, this is a common feature of \( \lor \lor \) proofs. Consider, for example, the proof of:

\[ P \lor Q, Q \lor R, \neg Q, \therefore P \land R \]

which proceeds as follows:

\[
\begin{align*}
(1) & \quad P \lor Q & \text{A} \\
(2) & \quad Q \lor R & \text{A} \\
(3) & \quad \neg Q & \text{A} \\
(4) & \quad \text{Show } (P \lor R) & \\
(5) & \quad P & \text{ACP} \\
(6) & \quad \text{Show } (Q \lor R) & \\
(7) & \quad Q & \text{ACP} \\
(8) & \quad Q & \text{ACP} \\
(9) & \quad \text{Show } P \land R & \\
(10) & \quad \neg P \land R & \text{AIP} \\
(11) & \quad Q & R, 8 \\
(12) & \quad \neg Q & R, 3 \\
(13) & \quad \text{Show } R \rightarrow P \land R & \\
(14) & \quad R & \text{ACP} \\
(15) & \quad P \land R & \&I, 6,14 \\
(16) & \quad P \land R & \lor E, 2,7,13 \\
(17) & \quad \text{Show } Q \rightarrow (P \lor R) & \\
(18) & \quad Q & \text{ACP} \\
(19) & \quad \text{Show } P \land R & \\
(20) & \quad \neg (P \lor R) & \text{ACP} \\
(21) & \quad Q & R, 18 \\
(22) & \quad \neg Q & R, 3 \\
(23) & \quad P \land R & \lor E, 1,5,17
\end{align*}
\]
Our desired conclusion of $P \& R$ (highlighted in blue above) shows up six times during this proof. How are we to know when we've gotten it enough times?

The short answer is that we've gotten it enough times when we've gotten it on the proof level that we want. Notice that the six occurrences of $P \& R$ in this proof all occur in different subproofs. At no point could we simply have reiterated or downward-transmitted a previous conclusion of $P \& R$. In this case, our eventual goal is to get $P \& R$ on the top level of the proof -- that is, the wholly unindented level. We don't succeed in doing this until the derivation of $P \& R$ on line 23 suffices to cancel the Show of line 4.

Let me emphasize that the proof we just did is a tricky one, and it shouldn't be surprising if you're not comfortable tackling one of this complexity yet. For now, I just want you to be able to see how each occurrence of $P \& R$ plays a role in its own subproof, and why the complex nested subproof structure of the proof requires us to derive $P \& R$ so many times.

Knowing when you've gotten your conclusion for the last time takes some experience with this sort of proof, but being very cautious with the subproof structure will help make the issue clearer.

**Question #3**: If I'm doing a proof starting with a disjunction in which both disjuncts are the same (like $P \lor P$), do I have to do two proofs of my conclusion, one from each of the two identical disjuncts?

**Answer**: No. As with other proof rules, you can cite the same line more than once as an input. So, for example, the proof of:

$P \lor P, \therefore P$

proceeds as follows:

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$P \lor P$</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>Show $P$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Show $P \rightarrow P$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$P$</td>
<td>ACP</td>
</tr>
<tr>
<td>5</td>
<td>$P \lor E$, 1,3,3</td>
<td></td>
</tr>
</tbody>
</table>

$\lor E$ here requires you to derive $P$ from each of the disjuncts in $P \lor P$. But since those two disjuncts are the same, you don't have to do it twice -- you can just appeal twice to the one derivation of $P \rightarrow P$.

**Question #4**: OK, what if I've got a disjunction with two different disjuncts, but the methods of proving the desired conclusion from each are the same, or overlap a great deal? Do I have to do all the work twice?

**Answer**: Unfortunately, yes. Consider the following proof of:

$P \lor (P \& Q) \lor ((P \& Q) \& R), \therefore Q$

<table>
<thead>
<tr>
<th>Line</th>
<th>Formula</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(P &amp; Q) \lor ((P &amp; Q) &amp; R)$</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>Show $Q$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Show $(P &amp; Q) \rightarrow Q$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$P &amp; Q$</td>
<td>ACP</td>
</tr>
<tr>
<td>5</td>
<td>$Q$</td>
<td>&amp;E, 4</td>
</tr>
<tr>
<td>6</td>
<td>Show $(P &amp; Q) &amp; R, Q) \rightarrow Q$</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$(P &amp; Q) &amp; R$</td>
<td>ACP</td>
</tr>
<tr>
<td>8</td>
<td>$P &amp; Q$</td>
<td>&amp;E, 7</td>
</tr>
<tr>
<td>9</td>
<td>$Q$</td>
<td>&amp;E, 9</td>
</tr>
<tr>
<td>10</td>
<td>$Q \lor E$, 1,3,6</td>
<td></td>
</tr>
</tbody>
</table>

Despite the fact that lines 8 and 9 are just a repetition of lines 4 and 5, they must be done again. Lines 4 and 5 are sealed within the subproof running from lines 3 to 5, and once that subproof is finished, that information cannot be exported sideways to the subproof running from lines 6 to 9. It's a bit of a nuisance, but at least it won't be hard to figure out what to do in the repeated proof steps.

**Question #5**: Help -- I'm completely baffled by $\lor E$ proofs. I never know how to get started, or what to do to finish one.

**Answer**: $\lor E$ proofs take some time to get used to. Try going back over the examples we've done so far, and work on seeing the overall structure of the proofs, ignoring the details for now. Schematically, $\lor E$ proofs typically look like this:
### Some disjunction

**Show if first disjunct, then something or other**

- **first disjunct**
  - ACP
  - Work, work, work
  - [finally, that something or other]

**Not done yet -- now the other half**

**Show if second disjunct, then something or other**

- **second disjunct**
  - ACP
  - Work, work, work
  - [again, that something or other]

-at last, the end -- [that something or other] \( \lor \) E

If you don't feel like you've done everything twice, then you probably haven't done your \( \lor \) E proof correctly.

### Some More Examples Using \( \lor \) E [Next]

We'll wrap up our discussion of \( \lor \) E by going through three more complex proofs, including some that combine \( \lor \) E with both conditional and indirect proof, and some that have nested levels of \( \lor \) E. As these examples will show, such proofs can get quite complicated. One of the keys to following them is to keep in mind their large structure (the way in which they nest levels of indirect proof, conditional proof, and disjunction elimination) before you try to follow the small details.

We'll start by proving an instance of the associativity of disjunction, by proving:

- \((P \lor Q) \lor R, \therefore P \lor (Q \lor R)\)

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Premise/Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((P \lor Q) \lor R) A</td>
</tr>
<tr>
<td>2</td>
<td>Show (P \lor (Q \lor R))</td>
</tr>
<tr>
<td>3</td>
<td>Show ((P \lor Q) \rightarrow (P \lor (Q \lor R)))</td>
</tr>
<tr>
<td>4</td>
<td>(P \lor Q) ACP</td>
</tr>
<tr>
<td>5</td>
<td>Show (P \rightarrow (P \lor (Q \lor R)))</td>
</tr>
<tr>
<td>6</td>
<td>P ACP</td>
</tr>
<tr>
<td>7</td>
<td>(P \lor (Q \lor R)) (\lor) I, 6</td>
</tr>
<tr>
<td>8</td>
<td>Show (Q \rightarrow (P \lor (Q \lor R)))</td>
</tr>
<tr>
<td>9</td>
<td>Q ACP</td>
</tr>
<tr>
<td>10</td>
<td>(Q \lor R) (\lor) I, 9</td>
</tr>
<tr>
<td>11</td>
<td>(P \lor (Q \lor R)) (\lor) I, 10</td>
</tr>
<tr>
<td>12</td>
<td>(P \lor (Q \lor R)) (\lor) I, 11</td>
</tr>
<tr>
<td>13</td>
<td>Show (R \rightarrow P \lor (Q \lor R))</td>
</tr>
<tr>
<td>14</td>
<td>R ACP</td>
</tr>
<tr>
<td>15</td>
<td>(Q \lor R) (\lor) I, 14</td>
</tr>
<tr>
<td>16</td>
<td>(P \lor (Q \lor R)) (\lor) I, 15</td>
</tr>
<tr>
<td>17</td>
<td>(P \lor (Q \lor R)) (\lor) E, 13, 13</td>
</tr>
</tbody>
</table>

Schematically, the structure of this proof is:

- **Conditional Proof**
  - **Dilemma Reasoning**
  - **Dilemma Reasoning**
  - **Conditional Proof**
  - **Conditional Proof**

Each instance of dilemma reasoning (one applied to the disjunction \((P \lor Q) \lor R\), and one applied to the disjunction \(P \lor Q\)) is accompanied by two conditional proofs. One of the conditional proofs of the larger dilemma reasoning itself contains a second dilemma reasoning, which then spawns two more conditional proofs.

**Question**: Why did we only need \(\lor I\) once to go from \(P\) to \(P \lor (Q \lor R)\), but twice to go either from \(Q\) to \(P \lor (Q \lor R)\) or from \(R\) to \(P \lor (Q \lor R)\)?

**Answer**: Because \(P \lor (Q \lor R)\) is a disjunction whose two disjuncts are \(P\), on the one hand, and \(Q \lor R\), on the other hand. Since \(\lor\) can add any disjunct we want, if we have \(P\) we can simply add \(Q \lor R\) to get \(P \lor (Q \lor R)\). If, however, we have \(Q\), we cannot add anything to get the desired disjunction, because we do not have either disjunct. Instead, we have to
proceed in two stages. First, we add $R$ to get $Q \lor R$. This is now the right-hand disjunct in our desired disjunction, so we can use $\lor$ again to add $P$ and get $P \lor (Q \lor R)$. The case starting with $R$ is similar.

Next we'll look at the proof of:

* $-P \lor (Q \land R), (R \lor -Q) \rightarrow (S \land T), (T \land U) \lor -(T \lor U). \therefore -P \lor U$

The proof proceeds as follows:

| (1) $-P \lor (Q \land R)$ | A |
| (2) $(R \lor -Q) \rightarrow (S \land T)$ | A |
| (3) $(T \land U) \lor -(T \lor U)$ | A |
| (4) Show $-P \lor U$ |
| (5) Show $-P \rightarrow (-P \lor U)$ |
| (6) $-P$ | ACP |
| (7) $-P \lor U$ | $\lor$ I, 6 |
| (8) Show $(Q \land R) \rightarrow (-P \lor U)$ |
| (9) $Q \land R$ | ACP |
| (10) $R$ | $\&E$, 9 |
| (11) $R \lor -Q$ | $\lor$ I, 10 |
| (12) $S \land T$ | $\rightarrow$ E, 2, 11 |
| (13) Show $(T \land U) \rightarrow U$ |
| (14) $T \land U$ | ACP |
| (15) $U$ | $\&E$, 14 |
| (16) Show $-(T \lor U) \rightarrow U$ |
| (17) $-(T \lor U)$ | ACP |
| (18) Show $U$ |
| (19) $-U$ | $\lor$ E, 3, 13, 16 |
| (20) $T$ | $\&E$, 14 |
| (21) $T \lor U$ | $\lor$ I, 20 |
| (22) $-(T \lor U)$ | $R$, 17 |
| (23) $U$ | $\lor$ I, 23 |
| (24) $-P \lor U$ | $\lor$ I, 1, 5, 8 |

Like the previous proof, this proof contains two instances of dilemma reasoning. Its large structure is:

Dilemma Reasoning
- Conditional Proof
- Dilemma Reasoning
- Conditional Proof
- Indirect Proof

For our last example, we will prove the result:

* $-((P \lor Q) \land (P \lor R)), \therefore -(P \lor (Q \land R))$

The proof proceeds as follows:

| (1) $-((P \lor Q) \land (P \lor R))$ | A |
| (2) Show $-((P \lor Q) \land (P \lor R))$ |
| (3) $P \lor (Q \land R)$ | AIP |
| (4) Show $P \rightarrow ((P \lor Q) \land (P \lor R))$ |
| (5) $P$ | ACP |
| (6) $P \lor Q$ | $\lor$ I, 5 |
| (7) $P \lor R$ | $\lor$ I, 5 |
| (8) $(P \lor Q) \land (P \lor R)$ | $\&I$, 6, 7 |
| (9) Show $(Q \land R) \rightarrow ((P \lor Q) \land (P \lor R))$ |
| (10) $Q \land R$ | ACP |
| (11) $Q$ | $\&E$, 10 |
| (12) $P \lor Q$ | $\lor$ I, 11 |
| (13) $R$ | $\&E$, 10 |
The large structure of this proof is:

**Indirect Proof**

**Dilemma Reasoning**

**Conditional Proof**

The Ninth Rule: Biconditional Elimination

We come now to the last of the five connectives of sentential logic -- the biconditional. You will probably be pleased to know that I've saved this one for last not because it's more difficult than the others, but rather because both of the rules for the biconditional are fairly trivial and raise no new issues.

Biconditionals, of course, are like conditionals running in both directions. Thus our rule of biconditional elimination ($\leftrightarrow E$) will be like our rule of conditional elimination, but running in both directions. Where the rule of $\rightarrow E$ allowed us to move from inputs of $\Phi \rightarrow \Theta$ and $\Phi$ to an output of $\Theta$, our rule of $\leftrightarrow E$ will allow us to move from an input of a biconditional $\Phi \leftrightarrow \Theta$ and either side of the biconditional to an output of the other side. The formal input-output specification of the rule will thus be:

\[
\begin{array}{c}
\Phi \leftrightarrow \Theta \\
\Phi \leftrightarrow \Theta \\
\Theta \\
\end{array}
\]

Let's look at a couple of applications of this new rule. First, we'll prove the following result:

- $P \rightarrow R$, $P \leftrightarrow Q$, $\therefore Q \rightarrow R$

This is an important result because it mirrors the following important principle:

- If two sentences are equivalent, then one can be substituted for the other as a premise of an argument without disturbing the validity of the argument.

In the example we're about to prove, $Q$ is substituted for $P$ on the basis of their equivalence, and we show that the truth of the biconditional is unaffected. The proof proceeds as follows:

\[
\begin{array}{l}
(1) P \rightarrow R \\
(2) P \leftrightarrow Q \\
(3) \text{Show } Q \rightarrow R \\
(4) Q \\
(5) P \\
(6) R \\
\end{array}
\]

Within the conditional proof, we use $\leftrightarrow E$ in the right-to-left direction to obtain $P$, which can then be combined with our given conditional to get $R$.

For our next example, let's look at one which combines $\leftrightarrow E$ with indirect proof. We'll prove the following result:

- $P \rightarrow (Q \& (P \rightarrow R))$, $Q \leftrightarrow \neg R$, $\therefore \neg P$

The proof proceeds as follows:

\[
\begin{array}{l}
(1) P \rightarrow (Q \& (P \rightarrow R)) \\
(2) Q \leftrightarrow \neg R \\
(3) \text{Show } \neg P \\
(4) P \\
(5) Q \& (P \rightarrow R) \\
(6) Q \\
(7) \neg R \\
(8) P \rightarrow R \\
(9) R \\
\end{array}
\]

Here $\leftrightarrow E$ is used in the left-to-right direction to pull $\neg R$ out of the biconditional $Q \leftrightarrow \neg R$, and thus derive half of the contradiction needed to complete the indirect proof.
Since $\leftrightarrow E$ is a very straightforward rule -- especially once $\rightarrow E$ has been mastered -- it doesn't give rise to a lot of questions. I'll just quickly go over a few possible points of confusion.

**Question #1**: You've mentioned each time whether you are using $\leftrightarrow E$ in the right-to-left or left-to-right direction. What does that mean, and do I need to mention it?

**Answer**: By the direction of use, I mean whether you have already the left-hand side of the biconditional, and are deriving the right-hand side (that's the left-to-right direction), or whether you have already the right-hand side of the biconditional, and are deriving the left-hand side (the left-to-right direction). Thus:

- **Left-to-right use of $\leftrightarrow E$**: $\Phi, \Phi \leftrightarrow \Theta, \therefore \Theta$
- **Right-to-left use of $\leftrightarrow E$**: $\Theta, \Phi \leftrightarrow \Theta, \therefore \Phi$

You don't need to indicate which direction you are using $\leftrightarrow E$ when you use it in proofs. I've mentioned it here just to make it easier to follow how I'm using the rule.

**Question #2**: From $\neg (P \leftrightarrow Q)$ and $P$, can I derive $\neg Q$?

**Answer**: No. I haven't mentioned this for a while, so let me emphasize once again: a proof rule can only be used on sentences that have the appropriate connective as their main connective. Since the main connective of $\neg (P \leftrightarrow Q)$ is a negation, rather than a biconditional, the rule of $\leftrightarrow E$ cannot take it as input.

**Question #3**: Can I use $\leftrightarrow E$ to derive either $P \rightarrow Q$ or $Q \rightarrow P$ from $P \leftrightarrow Q$?

**Answer**: Not directly. The rule of $\leftrightarrow E$, as we've set it up here, allows only one of the two sides of the biconditional as output, not one of the two directions of the biconditional. However, the conditionals in the two directions can easily be derived using $\leftrightarrow E$, so we can get the following result:

$\cdot P \leftrightarrow Q, \therefore (P \to Q) \land (Q \to P)$

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>$P \leftrightarrow Q$ A</td>
</tr>
<tr>
<td>(2)</td>
<td>Show $(P \to Q) \land (Q \to P)$</td>
</tr>
<tr>
<td>(3)</td>
<td>Show $P \to Q$</td>
</tr>
<tr>
<td>(4)</td>
<td>$P$ ACP</td>
</tr>
<tr>
<td>(5)</td>
<td>$Q$ ACP</td>
</tr>
<tr>
<td>(6)</td>
<td>Show $Q \to P$</td>
</tr>
<tr>
<td>(7)</td>
<td>$Q$ ACP</td>
</tr>
<tr>
<td>(8)</td>
<td>$P$ ACP</td>
</tr>
<tr>
<td>(9)</td>
<td>$(P \to Q) \land (Q \to P)$ &amp;I, 3,6</td>
</tr>
</tbody>
</table>

We could, had we wanted, have used (instead of our current rule) a rule of biconditional elimination which allowed us to extract conditionals from a biconditional; the final power of the proof system would have been unaffected by this change. As we'll see below, our rule of biconditional introduction does follow this model.

**Question #4**: If I have $P \leftrightarrow P$, can I use biconditional elimination on it?

**Answer**: Sure, as long as you've also got $P$. Of course, the output will be $P$, which you've already got, so it won't be terribly useful, but it will be valid. There's no rule against doing useless things in a proof.

**The Tenth Rule: Biconditional Introduction**

We come at last to the final proof rule of our system -- the rule of biconditional introduction. In setting out the rule of biconditional introduction, we suggested that the analogy between the conditional and the biconditional showed that we could take the $\rightarrow E$ inference pattern, make it bidirectional, and get an appropriate rule of $\leftrightarrow E$. If we were to follow the same idea with biconditional introduction, we would take the proof strategy of conditional proof and extend it to a new proof strategy of biconditional proof. This proof strategy would presumably involve, in effect, doing two conditional proofs - - one from an auxiliary assumption of $\Phi$ to a conclusion of $\Theta$, and one in the other direction from an auxiliary assumption of $\Theta$ to a conclusion of $\Phi$.

Introducing a new strategy of biconditional proof would be the most theoretically elegant way of providing a rule of biconditional elimination, but that's not what we're going to do. We've always had a tendency to treat the biconditional as somewhat of a second-class connective, thrown in just to avoid the need to keep writing conditionals in both directions.
We'll follow that tendency here, giving a rule of biconditional introduction which makes quite explicit the view that the biconditional is just an abbreviation for conditionals in both directions.

Our rule of biconditional introduction, then, will simply allow us to introduce a biconditional between $\Phi$ and $\Theta$ when we have conditionals in both directions between them. The formal input-output conditions for the rule are:

$$
\Phi \rightarrow \Theta
\longleftrightarrow
\Theta \rightarrow \Phi
\quad \text{--------}
\Phi \leftrightarrow \Theta
$$

We could have formulated a similar rule of $\leftrightarrow E$ which allowed us to go from $\Phi \leftrightarrow \Theta$ to either of the conditionals $\Phi \rightarrow \Theta$ and $\Theta \rightarrow \Phi$. However, since we would in most cases then want to use these conditionals with $\rightarrow E$ to extract their consequents, the version of $\leftrightarrow E$ we give above turns out usually to save time.

Let's look at a basic application of $\leftrightarrow I$. We'll prove the following result:

- $P \rightarrow Q$, $Q \rightarrow R$, $R \rightarrow P$, \[P \leftrightarrow R\]

The proof proceeds as follows:

1. $P \rightarrow Q$ A
2. $Q \rightarrow R$ A
3. $R \rightarrow P$ A
4. Show $P \leftrightarrow R$
5. Show $P \rightarrow R$
   - (6) $P$ ACP
   - (7) $Q$ $\rightarrow E$, 1,6
   - (8) $R$ $\rightarrow E$, 2,7
6. $P \leftrightarrow R$ $\leftrightarrow I$, 5,3

To prove $P \leftrightarrow R$ using $\leftrightarrow I$, we need to have the conditionals in both directions -- that is, both $P \rightarrow R$ and $R \rightarrow P$.

For another example, we'll prove:

- $P \rightarrow Q$, :: $P \leftrightarrow (P \& Q)$

The proof proceeds as follows:

1. $P \rightarrow Q$ A
2. Show $P \leftrightarrow (P \& Q)$
3. Show $P \rightarrow (P \& Q)$
   - (4) P ACP
   - (5) Q $\rightarrow E$, 1,4
4. Show $(P \& Q) \rightarrow P$
   - (7) $P \& Q$ ACP
   - (8) $P$ &E, 7
5. $P \leftrightarrow (P \& Q)$ $\leftrightarrow I$, 3,6

Again, to derive the biconditional we need conditionals in each direction. In this example, we aren't given either of the conditionals to begin with, so both must be derived using conditional introduction.

Now we'll look at two more complex examples. First, we'll prove the following result:

- $(P \leftrightarrow Q)$, :: $P \leftrightarrow \neg Q$

The proof proceeds as follows:

1. $(P \leftrightarrow Q)$ A
2. Show $P \leftrightarrow \neg Q$
3. Show $P \rightarrow \neg Q$
   - (4) P ACP
   - (5) Show $\neg Q$
   - (6) Q AIP
(7) Show $P \rightarrow Q$

| (8) $P$ | ACP |
| (9) $Q$ | R, 6 |

(10) Show $Q \rightarrow P$

| (11) $Q$ | ACP |
| (12) $P$ | R, 4 |

(13) $P \leftrightarrow Q$  \(\leftrightarrow I, 7,10\)

(14) $\neg(P \leftrightarrow Q)$  \(R, 1\)

(15) Show $\neg Q \rightarrow P$

| (16) $Q$ | ACP |

| (17) Show $P$

| (18) $\neg P$ | AIP |

(19) Show $P \rightarrow Q$

| (20) $P$ | ACP |

| (21) Show $Q$

| (22) $\neg Q$ | AIP |
| (23) $P$ | R, 20 |
| (24) $\neg P$ | R, 18 |

(25) Show $Q \rightarrow P$

| (26) $Q$ | ACP |

| (27) Show $P$

| (28) $\neg P$ | AIP |
| (29) $Q$ | R, 26 |
| (30) $\neg Q$ | R, 16 |

(31) $P \leftrightarrow Q$  \(\leftrightarrow I, 19,25\)

(32) $\neg(P \leftrightarrow Q)$  \(R, 1\)

(33) $P \leftrightarrow \neg Q$  \(\leftrightarrow I, 3,15\)

This proof is quite complicated. Here's a map of its large structure:

- **Biconditional Introduction**
- **Conditional Proof**
  - **Indirect Proof**
    - **Biconditional Introduction**
      - **Conditional Proof**
      - **Conditional Proof**
    - **Conditional Proof**
    - **Indirect Proof**
      - **Biconditional Introduction**
      - **Conditional Proof**
      - **Indirect Proof**
      - **Indirect Proof**

Here I'm using "Biconditional Introduction" to mark the beginning of the process of deriving a biconditional -- a process which is typically (but not always) followed by a pair of conditional proofs to derive the conditionals in both directions. The actual application of the rule of $\leftrightarrow I$, however, doesn't happen until the end of that process, after both conditional proofs have been completed.

It's much easier to follow this proof if you can see in it the large structure that I set out above. It's also easier to follow it if you can recognize certain inferential patterns that we've seen before, being applied here. So, for example, lines 19 through 24 repeat the process we've seen before of deriving the truth of a conditional from the falsity of the antecedent of that conditional. Lines 25 through 30 then use this process again. If you can recognize this little routine, then the strange sequence of reiterations that we find on lines 23 and 24, and again on lines 29 and 30, will be less baffling.

For our next example, we'll prove a result which is often tremendously useful in mathematical contexts. The result is:

- $P \rightarrow Q$, $Q \rightarrow R$, $R \rightarrow P$  \(\vdash(P \leftrightarrow Q) \& (Q \leftrightarrow R) \& (P \leftrightarrow R)\)

The proof proceeds as follows:

| (1) $P \rightarrow Q$  | A |

| (2) $Q \rightarrow R$  | A |

| (3) $R \rightarrow P$  | A |
This proof in essence performs our earlier proof of \(P \leftrightarrow R\) three times, and then combines the three resulting biconditionals. (Notice that I'm here allowing myself to use &I on three inputs at once to produce a trefoil conjunction. I'll allow this shortcut from now on.)

The result we just established – that from \(P \rightarrow Q\), \(Q \rightarrow R\), and \(R \rightarrow P\) you can derive \((P \leftrightarrow Q) \& (Q \leftrightarrow R) \& (P \leftrightarrow R)\) – has the following mathematical application. In mathematical contexts, it is frequently desirable to show that a number of conditions are all equivalent to one another. So, for example, it's a standard theorem in real analysis to prove that if \(S\) is a subset of \(R^3\), then the following three conditions are equivalent:

- \(S\) is closed and bounded.
- \(S\) is compact
- Every infinite subset of \(S\) has a limit point in \(S\).

(Don't worry too much if you don't have the math to know what this means – the main point is just that there are three conditions which are said to be equivalent.) Such theorems are extremely useful, because they allow one to freely interchange these conditions as one works on a proof. (We'll talk in a little bit about why this free interchange is possible.) But – as we just saw – proving that a bunch of sentences are all equivalent to each other takes a lot of work. For three sentences, three biconditionals (and hence six conditionals) have to be proved. For four sentences, six biconditionals (and hence 12 conditionals) have to be proved. For five sentences, 10 biconditionals (and hence 20 conditionals) have to be proved. The length of the proof quickly becomes prohibitive.

However, what the above proof shows us is that if we have three sentences, and the first implies the second, the second implies the third, and the third implies the first (and thus they form a circle of implication), then biconditionals connect all of them. Proving the three conditionals necessary to create a circle is an easier task than proving the six needed to establish explicit biconditionals in all directions. And the savings increase as the number of conditions increase – if there are six conditions, then only six conditionals are needed to create a circle of implication, but 30 conditionals have to be proved to get the explicit biconditionals. Note, by the way, that not any six conditionals will do – they must form a circle. If our six conditions are \(A, B, C, D, E,\) and \(F,\) then the collection of conditionals:

- \(A \rightarrow E, E \rightarrow C, C \rightarrow F, F \rightarrow D, D \rightarrow B, B \rightarrow A\)

forms a circle and thus gives us all the biconditionals, but the collection:

- \(A \rightarrow E, E \rightarrow A, C \rightarrow F, F \rightarrow D, D \rightarrow B, B \rightarrow C\)

does not form a circle, and leaves open the possibility that the pair \(A\) and \(E\) are not equivalent to the quartet \(B, C, D,\) and \(F.\)

Let's look now at a couple of questions about \(\leftrightarrow\):

**Question #1:** Biconditional introduction proofs always seem really long. Is there a way to avoid all that length?

**Answer:** No, not really. Biconditional introduction proofs are typically quite long because they are typically really two proofs in one. Since establishing the biconditional involves proving conditionals in both directions, you're really proving two results when you prove a biconditional. Sometimes you'll get lucky and be given one conditional direction for free, but often not. And if you're proving nested biconditionals, it gets even worse. A biconditional within a biconditional can often mean proving conditionals in four directions, making for quite a long proof.
Question #2: Does it matter what order I derive the two conditionals in? If I’m deriving $P \leftrightarrow Q$, do I need to get $P \rightarrow Q$ first and then $Q \rightarrow P$?

Answer: No, as with all rules with multiple inputs, the inputs can come from anywhere earlier in the proof (so long as there is no upward or sideways transmission of information).

Summary of the Proof System [Next]
We have now, at long last, finished setting out the entire proof system for sentential logic. In this section, we’ll give a very quick and condensed overview of the system so that you can have all the details in one place without slogging through all the exposition again.

A proof is a series of lines, each of which is either:
- An assertion line, containing some sentence of sentential logic and a proof rule (together with input lines) justifying that sentence
- A Show line, declaring an intention to derive some sentence, and invoking some proof strategy.

All of the proof rules are input-output filters, which take certain inputs (in the form of earlier lines in the proof) and produce certain outputs. The primary proof rules are introduction and elimination rules for each connective in the logic, and are specified as follows (in each case, the sentence(s) above the line give the input conditions, and the sentence below the line gives the output result):

<table>
<thead>
<tr>
<th></th>
<th>Elimination</th>
<th>Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conjunction</td>
<td>&amp;E: $\Phi &amp; \Theta$ or $\Theta &amp; \Phi$</td>
<td>&amp;I: $\Phi$ $\Theta$</td>
</tr>
<tr>
<td>Disjunction</td>
<td>vE: $\Phi \rightarrow \Psi$</td>
<td>vI: $\Phi$ $\Theta$</td>
</tr>
<tr>
<td>Conditional</td>
<td>$\Phi \rightarrow \Theta$</td>
<td>No rule. Use strategy of CONDITIONAL PROOF.</td>
</tr>
<tr>
<td>Biconditional</td>
<td>$\Phi \leftrightarrow \Theta$ or $\Theta \leftrightarrow \Phi$</td>
<td>$\Phi \rightarrow \Theta$</td>
</tr>
<tr>
<td>Negation</td>
<td>$\neg\neg\Phi$ $\Phi$</td>
<td>No rule. Use strategy of INDIRECT PROOF</td>
</tr>
</tbody>
</table>

Note: The double line between input and output on the rule of $\neg\neg$ indicates that this rule can run in either direction.

In addition to the primary rules, there are two structural rules:

A (Assumption): This rule allows an initial assumption to be recorded at the beginning of a proof. The A rule requires no inputs and can produce any output. It also has two variations for later in the proof:

ACP (Assumption for Conditional Proof): Allows an auxiliary assumption to be added to the beginning of a conditional proof.

AIP (Assumption for Indirect Proof): Allows an auxiliary assumption to be added to the beginning of an indirect proof.

These two variants can produce only certain outputs. See the specification of the proof strategies for the constraints on ACP and AIP.

R (Reiteration): This rule allows an earlier line of the proof to be repeated, so long as it does not violate the ban on downward and sideways transmission of information (see below).
The goal of a proof or a subproof is indicated using a Show line. Each Show line must invoke a proof strategy. A proof strategy determines (a) what is needed to complete the subproof invoked by the Show line, and (b) what auxiliary assumptions are given, if any. There are three available proof strategies (with two versions of one strategy):

**Direct Proof:**
- Can be used on: Show line of any form
- Auxiliary assumptions: None
- Success conditions: Derive sentence in Show line

**Conditional Proof:**
- Can be used on: Show line of form \( \Phi \rightarrow \Theta \)
- Auxiliary assumptions: \( \Phi \) (antecedent of conditional to be proved), using ACP
- Success conditions: Derive \( \Theta \) (consequent of conditional to be proved)

**Indirect Proof [standard version]:**
- Can be used on: Show line of form \( \neg \Phi \)
- Auxiliary assumptions: \( \Phi \) (unnegated form of sentence to be proved), using AIP
- Success conditions: Derive a contradiction

**Indirect Proof [alternative version]:**
- Can be used on: Show line of any form
- Auxiliary assumptions: Negated form of sentence to be shown, if that sentence is unnegated; unnegated form of sentence to be shown, if that sentence is negated
- Success conditions: Derive a contradiction

Every time a Show line is entered into a proof and a proof strategy is invoked, a subproof is initiated. The subproof concludes when the success conditions are met; at that point the Show is cancelled and the sentence on the Show line is treated like regular non-Show sentences. Subproofs can be nested; either indentation or bracketing is used to show the subproof structure of a proof.

Finally, there are constraints on what lines can validly serve as inputs to proof rules. The two constraints are that the input line must be asserted and available:

- A line is asserted if it is either a non-Show line occurring earlier in the proof, or a cancelled Show line. Thus uncancelled Show lines and (obviously) lines that haven't occurred yet are not valid inputs to proof rules.

- A line is available if it occurs either in the current subproof or in a higher-level proof of which the current subproof is a part. The requirement of availability is thus the same as the ban on upward or sideways transmission of information -- no information can be transmitted (i.e., no line can be used as input to a rule) from a subproof to a higher level proof or from a subproof to another subproof which is not nested within it.

A proof as a whole is successfully completed when its primary Show is cancelled. A successful proof demonstrates that the sentence on the primary Show line can be derived from whatever sentences (if any) appear via the A rule before the primary Show line.

**More Examples of Proofs [Next]**

Now that we've got the whole proof system in place, I want to go through five examples of fairly elaborate proofs, taking time to talk about the structure of each proof and how one might manage to find one's way through it. After going over these examples, we'll move on to talk some about strategic issues, and how -- once you've mastered all the proof rules -- you can get good at finding the right way to use the techniques the proof system offers you and construct successful proofs.

For our first example, we will prove the following result:
- \( (P \& Q) \leftrightarrow (P \& R) \; \therefore P \rightarrow (Q \leftrightarrow R) \)

After this and each of the other four proofs, I'll give a line-by-line commentary explaining for each important or tricky line why we made the move we did. The proof proceeds as follows:

| (1) (P \& Q) \leftrightarrow (P \& R) | A |
| (2) Show P \rightarrow (Q \leftrightarrow R) |
| (3) P | ACP |
| (4) Show Q \rightarrow R |
Here's the commentary:

**Line 2:** This is the major show line for the proof, and its content is dictated by what we are asked to prove. Since \( P \rightarrow (Q \leftrightarrow R) \) is a conditional, we invoke conditional proof. This dictates (a) that the next line gives us \( P \) by ACP, and (b) that we need \( Q \leftrightarrow R \) to complete the conditional proof and finish the whole proof.

**Line 4:** Our current goal for completing the proof is \( Q \leftrightarrow R \). We will get \( Q \leftrightarrow R \) using \( \leftrightarrow I \), which requires us to get conditionals in each direction. Here we start a conditional proof of \( Q \rightarrow R \), to give us one of the two conditionals we need.

**Line 6:** Here we combine the \( P \) we had from the ACP of line 3 (for the higher-level conditional proof) with the \( Q \) we have as the ACP for the lower-level conditional proof. The resulting \( P \& Q \) will be used as input to \( \leftrightarrow E \).

**Line 7:** Now that we have the left side of the biconditional, we can extract the right side with \( \leftrightarrow E \). The right side is useful because one of its conjuncts is \( R \), which is what we want.

**Line 8:** We finish the first conditional proof by extracting \( R \) from \( P \& R \).

**Line 9:** Next we start deriving the second conditional we need for the biconditional \( Q \leftrightarrow R \). Thus we start a conditional proof of \( R \rightarrow Q \). This second conditional proof is identical in structure to the first, but with the occurrences of \( Q \) and \( R \) reversed, so we'll go through the details rather briefly.

**Line 11:** We assemble the conjunction in preparation for \( \leftrightarrow E \).

**Line 12:** We use \( \leftrightarrow E \) in the right-to-left direction.

**Line 13:** We extract \( Q \) from \( P \& Q \), completing the conditional proof.

**Line 14:** Having deriving conditionals in both directions (on lines 4 and 9), we use \( \leftrightarrow I \) to assemble them to \( Q \leftrightarrow R \), which is what we need to complete the proof.

Here's a map of the large structure of the proof:

- **Biconditional Introduction**
- **Conditional Proof**
- **Conditional Proof**

For our second example, we'll prove the following result:

\[
(P \lor Q) \& T, \ Q \rightarrow ((R \& T) \rightarrow S), \ \therefore \ \neg P \rightarrow (R \rightarrow S)
\]

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>(1) ( P \lor Q ) &amp; T</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) ( Q \rightarrow ((R &amp; T) \rightarrow S) )</td>
<td>A</td>
</tr>
<tr>
<td>(3) Show ( \neg P \rightarrow (R \rightarrow S) )</td>
<td></td>
</tr>
<tr>
<td>(4) ( \neg P )</td>
<td>ACP</td>
</tr>
<tr>
<td>(5) Show ( R \rightarrow S )</td>
<td></td>
</tr>
<tr>
<td>(6) ( R )</td>
<td>ACP</td>
</tr>
<tr>
<td>(7) ( P \lor Q )</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(8) Show ( P \rightarrow Q )</td>
<td></td>
</tr>
<tr>
<td>(9) ( P )</td>
<td>ACP</td>
</tr>
<tr>
<td>(10) Show ( Q )</td>
<td></td>
</tr>
<tr>
<td>(11) ( \neg Q )</td>
<td>AIP</td>
</tr>
<tr>
<td>(12) ( P )</td>
<td>R, 9</td>
</tr>
<tr>
<td>(13) ( \neg P )</td>
<td>R, 4</td>
</tr>
<tr>
<td>(14) Show ( Q \rightarrow Q )</td>
<td></td>
</tr>
<tr>
<td>(15) ( Q )</td>
<td>ACP</td>
</tr>
<tr>
<td>(16) ( Q )</td>
<td>&amp;E, 1</td>
</tr>
<tr>
<td>(17) ( R &amp; T \rightarrow S )</td>
<td>&amp;E, 2, 16</td>
</tr>
<tr>
<td>(18) ( T )</td>
<td>&amp;E, 1</td>
</tr>
</tbody>
</table>
Here is a line-by-line commentary on the proof:

Line 3: This is the main Show line for the proof, and once it is cancelled, we will be done. Since we’re trying to show $\neg P \rightarrow (R \rightarrow S)$, we will use conditional proof, and our success condition will be deriving $R \rightarrow S$.

Line 5: Since our success condition -- $R \rightarrow S$ -- is itself a conditional, we invoke conditional proof once again. This is a frequent pattern when nested conditionals are involved in a proof. Our success condition now becomes deriving $S$. The obvious way to do this is via the nested conditional $Q \rightarrow ((R \& T) \rightarrow S)$, which is our second assumption.

To get $S$ out of this, we need both $Q$ and $R \& T$, and then two applications of $\rightarrow E$. We’ll start by trying to derive $Q$.

Line 7: This is probably the trickiest move in the proof. We’re trying to derive $Q$ to work on extracting $S$ from our second premise, and we extract $P \lor Q$ from our first premise because we’re going to use it in a dilemma reasoning to derive $Q$. (The key to seeing why this is a good route to getting $Q$ is seeing that we’re already halfway to $Q$ when we have $P \lor Q$ -- all that remains, in effect, is to get rid of the $P$ option.)

Line 8: Now we begin the first half of the dilemma reasoning, deriving the conditional $P \rightarrow Q$. This derivation will follow a pattern we’ve seen before -- we already know that $\neg P$, so the false antecedent will guarantee the true conditional.

Line 10: To derive $P \rightarrow Q$, we need $Q$. There’s no obvious way to get $Q$ (after all, $Q$ is what we’re after in the higher-level dilemma reasoning as well), so we’ll use an indirect proof.

Line 11: We introduce the $\neg Q$ as AIP, but this assumption isn’t actually going to play a role in the indirect proof.

Line 12: We reiterate $P$ from line 9, giving us one half of our contradiction.

Line 13: We reiterate $\neg P$ from line 4, giving us the other half of our contradiction and completing the indirect proof of $Q$ and thereby also the conditional proof of $P \rightarrow Q$.

Line 14: Having derived $P \rightarrow Q$, we just need $Q \rightarrow Q$ to wrap up the dilemma reasoning. $Q \rightarrow Q$, of course, is very easy to derive using conditional proof, which is what we start doing on this line.

Line 15: Here we quickly end the conditional proof of $Q \rightarrow Q$ by assuming $Q$ by ACP. Since $Q$ is also what we need to complete the conditional proof, we’re done.

Line 16: Now that we’ve gotten both $P \rightarrow Q$ and $Q \rightarrow Q$, we can combine these two conditionals with the fact (from line 7) that $P \lor Q$, and conclude $Q$. At this point, the difficult part of the proof is done; there remains only some wrapping-up moves in which we use what we’ve derived so far.

Line 17: We use the $Q$ we’ve derived to strip off the first conditional level of $Q \rightarrow ((R \& T) \rightarrow S)$.

Line 18: We now need $R \& T$, which we’ll get by using $\& I$ on $R$ and on $T$. Here we pull $T$ out of the conjunction which was our first premise.

Line 19: We now combine $R \& T$ (from line 6) with the $T$ we just derived to get the final antecedent we need.

Line 20: A final application of $\rightarrow E$ gets $S$, and completes the proof. Notice that $S$ completes the conditional proof of $R \rightarrow S$ and cancels the Show of line 5, and that cancellation makes $R \rightarrow S$ available and completes the conditional proof of $\neg P \rightarrow (R \rightarrow S)$. It’s not uncommon for two subproofs to conclude simultaneously in this manner.

Here is a map of the large structure of the proof:

- Conditional Proof
  - Conditional Proof
  - Dilemma Reasoning
    - Conditional Proof
    - Indirect Proof
    - Conditional Proof

For our third example, we’ll prove the following result:

- $P \& Q \therefore P \leftrightarrow (Q \leftrightarrow (P \lor Q))$

The proof proceeds as follows:

| (1) $P \& Q$ | A |
| (2) Show $P \leftrightarrow (Q \leftrightarrow (P \lor Q))$ |
| (3) Show $P \rightarrow (Q \leftrightarrow (P \lor Q))$ |
| (4) $P$ | ACP |
| (5) Show $Q \rightarrow (P \lor Q)$ |
| (6) $Q$ | ACP |
| (7) $P \lor Q$ | $\lor I$, 6 |
| (8) Show $(P \lor Q) \rightarrow Q$ |
| (9) $P \lor Q$ | ACP |
| (10) $Q$ | $\& I$, 1 |
| (11) $Q \leftrightarrow (P \lor Q)$ | $\leftrightarrow I$, 5,8 |
Here is a line-by-line commentary on the proof:

**Line 3:** We're trying eventually to show a biconditional — \(P \leftrightarrow (Q \leftrightarrow (P \lor Q))\) -- to complete the Show of line 2 (which triggers a direct proof). To get this biconditional, we will derive conditionals in both directions and then apply \(\leftrightarrow I\).

**Line 4:** This line begins a conditional proof to derive the conditional in the left-to-right direction. To succeed in this conditional proof, we need to derive \(Q \leftrightarrow (P \lor Q)\).

**Line 5:** Our success condition for the conditional proof is \(Q \leftrightarrow (P \lor Q)\). Since this is a biconditional, to derive it we need again to prove conditionals in both directions. Here we begin proving the conditional in the left-to-right direction.

**Line 6:** We complete the conditional proof of \(Q \rightarrow (P \lor Q)\) by using \(\lor I\) to make the small inferential step from \(Q\) to \(P \lor Q\).

**Line 7:** We now tackle the right-to-left direction of the attempt to derive \(Q \leftrightarrow (P \lor Q)\). Here we assume \(P \lor Q\), and attempt to derive \(Q\).

**Line 8:** We get \(Q\), as it turns out, not by using the ACP at all, but by simply extracting it from our starting conjunction.

**Line 9:** Having gotten conditionals in two directions, we put them together with \(\leftrightarrow I\) to form \(Q \leftrightarrow (P \lor Q)\). This completes our first large conditional proof, and puts us halfway to the derivation of the final target biconditional \(P \leftrightarrow (Q \leftrightarrow (P \lor Q))\).

**Line 10:** Next we derive the conditional in the right-to-left direction. This direction, as it turns out, will be much simpler than the left-to-right direction.

**Line 11:** We ignore the rather complex ACP of \(Q \leftrightarrow (P \lor Q)\), which it turns out we don't actually need, and directly infer \(P\) from our starting assumption \(P \land Q\). This completes the right-to-left direction of our desired biconditional.

**Line 12:** We now piece together the two conditionals from lines 5 and 12 using \(\leftrightarrow I\) to derive \(P \leftrightarrow (Q \leftrightarrow (P \lor Q))\), and thus complete the proof.

Here is a map of the large structure of the proof:

- **Biconditional Introduction**
- **Conditional Proof**
- **Biconditional Introduction**
- **Conditional Proof**
- **Conditional Proof**
- **Conditional Proof**

It's a good (and challenging) exercise to show that you can also get a proof going in the opposite direction -- starting from \(P \leftrightarrow (Q \leftrightarrow (P \lor Q))\), and deriving \(P \land Q\).

For our fourth example, we'll prove the following result:

\[
\star \quad (P \rightarrow Q) \lor (Q \rightarrow P)
\]

The proof proceeds as follows:

| (1) Show \((P \rightarrow Q) \lor (Q \rightarrow P)\) |
| (2) Show \(P \lor \lnot P\) |
| (3) \(\lnot (P \lor \lnot P)\)  AIP |
| (4) Show \(\lnot P\) |
| (5) \(P\) AIP |
| (6) \(P \lor \lnot P\)  \(\lor I, 5\) |
| (7) \(\lnot (P \lor \lnot P)\)  R, 3 |
| (8) \(P \lor \lnot P\)  \(\lor I, 4\) |
| (9) Show \(P \rightarrow ((P \rightarrow Q) \lor (Q \rightarrow P))\) |
| (10) \(P\) ACP |
| (11) Show \(Q \rightarrow P\) |
| (12) \(Q\) ACP |
| (13) \(P\)  R, 10 |
| (14) \((P \rightarrow Q) \lor (Q \rightarrow P)\)  \(\lor I, 11\) |
| (15) Show \(\lnot P \rightarrow ((P \rightarrow Q) \lor (Q \rightarrow P))\) |
| (16) \(\lnot P\) ACP |
| (17) Show \(P \rightarrow Q\) |
| (18) \(P\) ACP |
| (19) Show \(Q\) |
Here we do the work of deriving lines 3 - 8. Here is a map of the large structure of the proof:

### Here is a line-by-line commentary on the proof:

**Line 2:** The second line is the bold move of the proof, and is not something that would be likely to occur to someone working through the proof for the first time. We apparently just set aside entirely the goal of proving $(P \rightarrow Q) \lor (Q \rightarrow P)$, and start to work instead on proving $P \lor \neg P$. We do this because disjunctive conclusions are typically easiest to prove using disjunction assumptions (together with dilemma reasoning), but here we have no disjunctive assumption, so we're going to create one. Since $P \lor \neg P$ is a tautology, we know we can derive it for free, and then use its disjuncts -- from $P$ we can derive $Q \rightarrow P$, and from $\neg P$ we can derive $P \rightarrow Q$.

**Lines 3 - 8:** Here we do the work of deriving $P \lor \neg P$. We've seen this proof before, so I won't go over the details. Just notice that we can take a method of proving $P \lor \neg P$ and simply insert it into the current proof -- this is a consequence of the modular structure of the proof system.

**Line 9:** Now that we've derived $P \lor \neg P$, we will use it to trigger a dilemma reasoning. We thus need to derive two conditionals, one from $P$ to $(P \rightarrow Q) \lor (Q \rightarrow P))$, and one from $\neg P$ to $(P \rightarrow Q) \lor (Q \rightarrow P))$. Each of these two conditional proofs will derive one of the two conditionals in the desired disjunction, and then use $\lor I$ to get the disjunction. In our first conditional proof, which starts on this line, we will be proving from $P$, and will derive $Q \rightarrow P$.

**Line 11:** Here we declare our intention to show $Q \rightarrow P$ using conditional proof. This will be quite easy since we already know that $P$ is true, and $P$ is the goal of the conditional proof.

**Line 13:** Here we reiterate $P$, as given as ACP on line 10, and thereby complete the conditional proof of $Q \rightarrow P$.

**Line 14:** Having derived $Q \rightarrow P$, we use $\lor I$ to stick on a left disjunct and get $(P \rightarrow Q) \lor (Q \rightarrow P)$. This gives us one of the two conditionals we need for our dilemma reasoning.

**Line 15:** We now begin the other branch of the dilemma reasoning, this time reasoning from the other half of our disjunction -- $\neg P$. From $\neg P$ we will derive $P \rightarrow Q$.

**Line 17:** Here we begin a conditional proof of $P \rightarrow Q$.

**Line 19:** To derive $Q$, we use indirect proof, in which we will take advantage of the fact that we've already derived $P$ and $\neg P$.

**Line 21:** We reiterate $P$ from the conditional proof of $P \rightarrow Q$, giving us the first half of our contradiction.

**Line 22:** We reiterate $\neg P$ from the conditional proof of $\neg P \rightarrow ((P \rightarrow Q) \lor (Q \rightarrow P))$, giving us the second half of our contradiction, and completing the indirect proof (and thereby also the conditional proof of $P \rightarrow Q$).

**Line 23:** Having derived $P \rightarrow Q$, we use $\lor I$ to add a right disjunct and get $(P \rightarrow Q) \lor (Q \rightarrow P)$. This completes the proof of second conditional we need for the dilemma reasoning, and sets us up to end the proof.

**Line 24:** We combine the disjunction $P \lor \neg P$ which we derived in lines 2 through 8 with the two conditionals we then derived, and use $\lor E$ to get our final conclusion of $(P \rightarrow Q) \lor (Q \rightarrow P)$.

Here is a map of the large structure of the proof:

- Dilemma Reasoning
  - Indirect Proof
  - Conditional Proof
- Conditional Proof
  - Indirect Proof

For our last example, we will prove the following result:

- $(P \rightarrow Q) \lor (R \rightarrow S) \lor (P \rightarrow S) \lor (R \rightarrow Q)$

Notice that if we are thinking of the conditional as expressing implication, this result is a very odd one, since it would assert that, given any two claims, either the first implies the second or the second implies the first. However, this is manifestly untrue, as we can see from examples like:

- The moon is made of green cheese.
- Cary Grant is a talented actor.

neither of which implies the other. If, on the other hand, we think of the conditionals as simply expressing the material conditional, then the result is unproblematic. $P$ must be either true or false. If it is true, then $Q \rightarrow P$ is true, because it has a true consequent. If it is false, then $P \rightarrow Q$ is true, because it has a false antecedent. This shows that we must be cautious in taking theorems involving conditionals as showing something about the structure of implication.

For our last example, we will prove the following result:

- $(P \rightarrow Q) \lor (R \rightarrow S) \lor (P \rightarrow S) \lor (R \rightarrow Q)$
The proof proceeds as follows:

<table>
<thead>
<tr>
<th>Line</th>
<th>Statement</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((P \rightarrow Q) \lor (R \rightarrow S))</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>Show ((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Show ((P \rightarrow Q) \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)))</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(P \rightarrow Q)</td>
<td>ACP</td>
</tr>
<tr>
<td>5</td>
<td>Show (!P \lor Q)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(!P \lor Q)</td>
<td>AIP</td>
</tr>
<tr>
<td>7</td>
<td>Show (!P)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>(P)</td>
<td>AIP</td>
</tr>
<tr>
<td>9</td>
<td>(Q)</td>
<td>(\rightarrow), 4, 8</td>
</tr>
<tr>
<td>10</td>
<td>(!P \lor Q)</td>
<td>(\lor), 9</td>
</tr>
<tr>
<td>11</td>
<td>(!P \lor Q)</td>
<td>(\lor), 6</td>
</tr>
<tr>
<td>12</td>
<td>(!P \lor Q)</td>
<td>(\lor), 7</td>
</tr>
<tr>
<td>13</td>
<td>Show (!P \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)))</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>(!P)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Show (P \rightarrow S)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>(P)</td>
<td>ACP</td>
</tr>
<tr>
<td>17</td>
<td>Show (S)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>(!S)</td>
<td>AIP</td>
</tr>
<tr>
<td>19</td>
<td>(P)</td>
<td>(\rightarrow), 16</td>
</tr>
<tr>
<td>20</td>
<td>(!P)</td>
<td>(\rightarrow), 14</td>
</tr>
<tr>
<td>21</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\lor), 15</td>
</tr>
<tr>
<td>22</td>
<td>Show ((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>ACP</td>
</tr>
<tr>
<td>24</td>
<td>Show (R \rightarrow Q)</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>(R)</td>
<td>ACP</td>
</tr>
<tr>
<td>26</td>
<td>(Q)</td>
<td>(\rightarrow), 23</td>
</tr>
<tr>
<td>27</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\lor), 24</td>
</tr>
<tr>
<td>28</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\rightarrow), 12, 13, 22</td>
</tr>
<tr>
<td>29</td>
<td>Show ((R \rightarrow S) \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)))</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>((R \rightarrow S) \lor (R \rightarrow Q))</td>
<td>ACP</td>
</tr>
<tr>
<td>31</td>
<td>Show (!R \lor S)</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>(!(!R \lor S))</td>
<td>AIP</td>
</tr>
<tr>
<td>33</td>
<td>Show (!R)</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>(!R)</td>
<td>AIP</td>
</tr>
<tr>
<td>35</td>
<td>(!S)</td>
<td>(\rightarrow), 30, 34</td>
</tr>
<tr>
<td>36</td>
<td>(!R \lor S)</td>
<td>(\lor), 35</td>
</tr>
<tr>
<td>37</td>
<td>(!(!R \lor S))</td>
<td>(\lor), 32</td>
</tr>
<tr>
<td>38</td>
<td>(!R \lor S)</td>
<td>(\lor), 33</td>
</tr>
<tr>
<td>39</td>
<td>Show (!R \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)))</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>(!R)</td>
<td>ACP</td>
</tr>
<tr>
<td>41</td>
<td>Show (R \rightarrow Q)</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>(!R)</td>
<td>ACP</td>
</tr>
<tr>
<td>43</td>
<td>Show (Q)</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>(!Q)</td>
<td>AIP</td>
</tr>
<tr>
<td>45</td>
<td>(!R)</td>
<td>(\rightarrow), 42</td>
</tr>
<tr>
<td>46</td>
<td>(!R)</td>
<td>(\rightarrow), 40</td>
</tr>
<tr>
<td>47</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\lor), 41</td>
</tr>
<tr>
<td>48</td>
<td>Show ((S \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)))</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>(S)</td>
<td>ACP</td>
</tr>
<tr>
<td>50</td>
<td>Show (P \rightarrow S)</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>(!P)</td>
<td>ACP</td>
</tr>
<tr>
<td>52</td>
<td>(!S)</td>
<td>(\rightarrow), 49</td>
</tr>
<tr>
<td>53</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\lor), 50</td>
</tr>
<tr>
<td>54</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\rightarrow), 38, 39, 48</td>
</tr>
<tr>
<td>55</td>
<td>((P \rightarrow S) \lor (R \rightarrow Q))</td>
<td>(\rightarrow), 1, 3, 29</td>
</tr>
</tbody>
</table>
Here is a line-by-line commentary on the above proof:

**Line 3:** This Show line triggers a conditional proof. We want to show this conditional to get one of the two conditionals that we'll need to derive our final conclusion from the disjunctive assumption on line 1, using \( \lor E \).

**Line 5:** This is one of the least obvious steps in the proof. We've got the conditional \( P \rightarrow Q \), and we want to prove from it the disjunction \( (P \rightarrow S) \lor (R \rightarrow Q) \). In general, it's easiest to reach disjunctive conclusions from disjunctive premises, since one can then use one disjunct of the premise to derive one disjunct of the conclusion, and the other disjunct of the premise to derive the other disjunct of the conclusion. Since we know already that \( P \rightarrow Q \) is equivalent to \( \neg P \lor Q \), we're going to try to find a way in the proof system to convert the conditional to the disjunction, on the assumption that the disjunction will be more useful in the proof. The Show on this line invokes the strategy of indirect proof, so in the next line we enter \( \neg (P \lor Q) \), and start trying to derive a contradiction.

**Line 7:** This is another less-than-obvious line. We're trying to reach a contradiction to complete the indirect proof begun on line 5. Our AIP is \( \neg (\neg P \lor Q) \), so one possible source of a contradiction would be \( \neg P \lor Q \). If we could get \( \neg P \), we could then use \( \lor L \) to get \( \neg P \lor Q \). It's pretty much impossible to see ahead of time that this is going to work -- you just have to try out various options for getting the contradiction until something works. The Show on this line also invokes an indirect proof, so now we're two levels deep in indirect proof.

**Line 9:** Here we combine the P we got for AIP with the conditional \( P \rightarrow Q \) we have on line 4. It turns out that this is going to do useful things for us, but even if you don't see that ahead of time, you should be tempted to make this move just on the general principle that conditionals and antecedents should be combined using \( \rightarrow E \) whenever possible.

**Line 10:** Another devious move -- we take the Q we just derived and use \( \lor I \) on it to make \( \neg P \lor Q \), which will serve as half of our contradiction (for the lower-level indirect proof). \( \lor I \) moves are often tricky points in proofs, because there are so many choices for what to take as the new disjunct.

**Line 11:** Here we reiterate the \( \neg (P \lor Q) \) we had as our AIP for the higher-level indirect proof, and thus get the second half of the contradiction for our lower-level indirect proof. Note that this contradiction brings to a close the lower-level indirect proof started on line 7, but doesn't yet end the higher-level indirect proof started on line 5. We're almost there, though...

**Line 12:** Once we obtain \( \neg P \) by finishing the lower-level indirect proof, we can then use it with \( \lor L \) to get \( \neg P \lor Q \) (again, but on a higher level) to contradict our AIP \( \neg (P \lor Q) \). Note that this sequence is made extra-confusing by the fact that we use the very same contradiction (between \( \neg P \lor Q \) and \( \neg (\neg P \lor Q) \)) in both the higher-level and lower-level indirect proofs. Make sure you are perfectly clear on why we need to derive this same contradiction twice.

**Line 13:** Now that we've converted \( P \rightarrow Q \) into the disjunction \( \neg P \lor Q \), we begin the process of using dilemma reasoning to get \( (P \rightarrow S) \lor (R \rightarrow Q) \) out of each disjunct. The grand strategy will be to show that one disjunct -- \( \neg P \) -- allows us to derive one disjunct -- \( P \rightarrow S \) -- and the other disjunct -- \( Q \) -- allows us to derive the other disjunct -- \( R \rightarrow Q \). So from here we will perform two conditional proofs, which will be followed by an \( \lor E \) on line 28. On this line we begin the first conditional proof.

**Line 15:** Our goal is to derive \( (P \rightarrow S) \lor (R \rightarrow Q) \) from \( \neg P \), and we will reach this goal by deriving \( P \rightarrow S \) and then using \( \lor I \). So on this line we begin a conditional proof of \( P \rightarrow S \). This conditional proof will follow the pattern we have seen before of deriving the truth of a conditional from the falsity of the antecedent -- we will use indirect proof to get \( S \), and then reiterate \( P \) and \( \rightarrow P \) from earlier.

**Line 17:** Here we begin the indirect proof of \( S \), which we need to complete the conditional proof of \( P \rightarrow S \).

**Line 18:** \( \neg S \) is taken as AIP here, but it won't actually be used in the proof (since we already have the contradiction we need).

**Line 19:** We reiterate \( P \), which we had as the ACP for proving \( P \rightarrow Q \).

**Line 20:** We reiterate \( \neg P \), which we had as the ACP for proving \( \rightarrow P \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)) \). \( \neg P \) is thus transmitted down through two subproof levels. This gives us our contradiction and both (a) completes the indirect proof of \( S \) and thereby (b) completes the conditional proof of \( P \rightarrow S \).

**Line 21:** Having derived \( P \rightarrow S \), we use \( \lor I \) to add a right disjunct and get \( (P \rightarrow S) \lor (R \rightarrow Q) \). This completes the conditional proof started on line 15.

**Line 22:** Here we begin the other half of the dilemma reasoning from \( \neg P \lor Q \) to \( (P \rightarrow S) \lor (R \rightarrow Q) \). This half consists of a conditional of \( Q \rightarrow ((P \rightarrow S) \lor (R \rightarrow Q)) \). As before, we will prove only one half of this disjunction from \( Q \), and then use \( \lor I \). In this case, we will prove \( R \rightarrow Q \).

**Line 24:** Here we begin a conditional proof of \( R \rightarrow Q \), to fulfill the plan we set out in the notes for line 22. This conditional proof will be quite trivial, since the success conditions for it will be \( Q \), which we already have.

**Line 26:** Here we reiterate \( Q \) from line 23, thus completing the conditional proof.

**Line 27:** Next we use \( \lor I \) to add a left disjunct to \( R \rightarrow Q \) and get \( (P \rightarrow S) \lor (R \rightarrow Q) \), which completes the conditional proof started on line 22.

**Line 28:** Having derived both of the conditionals we need for the dilemma reasoning, we now apply \( \lor E \) to derive \( (P \rightarrow S) \lor (R \rightarrow Q) \) directly from the disjunction \( \neg P \lor Q \). This completes the conditional proof started on line 3, and brings us to the halfway point in the proof. We now have one of the two conditionals we need for the dilemma reasoning which forms the main structure of the proof.
Here is a map of the large structure of the proof:

Dilemma Reasoning

Conditional Proof
Indirect Proof

Indirect Proof

Dilemma Reasoning
Conditional Proof
Conditional Proof

Conditional Proof

Indirect Proof

Dilemma Reasoning
Conditional Proof

Conditional Proof

Indirect Proof

Common Mistakes in Using the Proof System [Next]

At this point you have seen the entire proof system for sentential logic, and had a chance to look over many proofs in that system. Now it's time for you to start developing your own skill at producing proofs in the system. Looking over proofs that have already been done can make the process look deceptively easy -- you can end up understanding all the moves that I make, and thinking that you've got things under control, but then discover when you are set off on your own to produce a proof that you don't really know how to go about doing it.

We will cover two types of advice for successful proof-craft. The first, which we'll cover in this section, is a list of common mistakes that people make in the mechanics of using the proof rules, and some thoughts on how to avoid those mistakes. The second is a collection of strategic hints to help give guidance when you're not sure what to do next in a proof.

One of the most common difficulties beginners at proofs have is that they simply apply the proof rules incorrectly -- using them in a situation in which the necessary inputs are not available, or using them with the wrong kinds of inputs, or drawing with them conclusions not licensed by the rule. The temptation to misapply proof rules can be great. Quite often one knows what one needs to derive to finish a proof, and even knows which sentences will probably lead to deriving that something, but can't quite see how to make it work. In this situation, it can be easy to just force a proof rule to work by slightly abusing it (especially if one is not quite clear on how the rules work to begin with). Suppose, for example, you're trying to prove the following result:

\[ P \lor Q, P \rightarrow R, R \rightarrow Q \therefore \neg R \]

You start the proof like this:
Looking back at your assumptions, you see that $P \rightarrow \neg R$ will give you what you want, if you can only get $P$. But how to get $P$? There’s no immediately obvious correct way to do this, so if you’re sufficiently convinced that getting $P$ must be the key to completing the proof, you might convince yourself that you must be able to use $\lor E$ to get $P$, like this:

| (1) $P \lor Q$ | A |
| (2) $P \rightarrow \neg R$ | A |
| (3) $R \rightarrow \neg Q$ | A |
| (4) Show $\neg R$ | A |
| (5) $P \lor E, 1$ | A |
| (6) $\neg R \rightarrow E, 2, 5$ | A |

Of course, this proof is incorrect because $\lor E$ does not allow you to infer one disjunct from a disjunction.

The first thing to realize about misusing proof rules is that it is a very serious error in constructing proofs. Misusing a proof rule is like making an illegal move in chess. Obtaining a checkmate by making an illegal move does nothing to show your chess skills, and completing a proof by misusing a rule does nothing to show your proof-construction skills. Anyone can win by cheating; there’s no virtue in it. And “completing” a proof by misusing a proof rule often disguises the real complexities of a proof. One might think that the abuse of $\lor E$ in the example above, even if it is wrong, is only a small error. However, that error causes one to miss entirely the interesting part of the proof. Here’s the proof done properly:

| (1) $P \lor Q$ | A |
| (2) $P \rightarrow \neg R$ | A |
| (3) $R \rightarrow \neg Q$ | A |
| (4) Show $\neg R$ | A |
| (5) Show $P \rightarrow \neg R$ | A |
| (6) $P \rightarrow E, 2, 6$ | A |
| (7) $\neg R \rightarrow E, 2, 6$ | A |
| (8) Show $Q \rightarrow \neg R$ | A |
| (9) $Q \rightarrow E, 2, 6$ | A |
| (10) Show $\neg R$ | A |
| (11) $R \rightarrow E, 3, 12$ | A |
| (12) $\neg Q \rightarrow AIP$ | A |
| (13) $Q \rightarrow R, 9$ | A |
| (14) $\neg R \lor E, 1, 5, 8$ | A |

The correct proof contains a dilemma reasoning within which there are two conditional proofs, one of which contains an indirect proof. None of this comes out in the improper proof abusing $\lor E$. That improper proof is akin to a game of chess in which white moves his queen and rook together to checkmate on the first move -- you don’t really learn much about chess strategy from seeing that happen.

So if you’re going to get good at constructing proofs, it’s absolutely essential that you learn to use all the proof rules correctly. You simply won’t be able to make any progress on strategic issues until you’ve got the mechanics down. In a minute I will go over the most common errors with the proof rules, but before doing this, I want to give one simple tip that can avoid many of the problems.

Since each of the proof rules is an input-output filter, each rule can be associated with a characteristic number -- the number of inputs which that rule requires (all of the rules produce exactly one output, so there’s no variation there). Here’s a quick summary of the proof rules categorized by the number of inputs they require:

<table>
<thead>
<tr>
<th>0 inputs</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 input</td>
<td>R</td>
</tr>
<tr>
<td>&amp;E</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>AIP</td>
</tr>
<tr>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>\lor</td>
</tr>
</tbody>
</table>
If you are having trouble getting the application of proof rules down, I strongly recommend as a first step simply checking to make sure that you’re using the right number of inputs for each rule. Make sure that when you use $\lor E$, you provide three inputs, and that you never provide three inputs to any other rule. Make sure that $\land I$, $\rightarrow E$, $\leftrightarrow E$, and $\leftrightarrow I$ all receive two inputs. Provide just one input, not two or three (or zero) to the rules $R$, $\&E$, $\lor I$, and $\rightarrow E$. And finally, don’t provide any inputs to the assumption rules $A$, $ACP$, and $AI\neg P$.

Most of the ways of misunderstanding rules involve getting the structure of the rules wrong so as to change the number of inputs, so this quick numerical scan will catch many of the worst errors. Of course, it won’t catch everything (it won’t, for example, catch it if you somehow use $\rightarrow E$ in the way that $\land I$ should be used), but it’s a good start.

Now let’s go through specific problems with specific rules. I won’t discuss all the rules here; just those that are most likely to cause difficulties.

**Problems with $\lor E$**

This is the rule that is most likely to give beginners difficulty. Since $\lor E$ is a rather complicated rule, it can become quite easy to give in to the temptation to simplify it. The most frequent error is to use $\lor E$ on analogy with $\&E$, producing “proof” like the following:

1. $P \lor Q$ A
2. Show $Q$
3. $Q \lor E, 1$

This proof is incorrect. $\lor E$ does not allow you to infer a disjunct directly from a disjunction. Remember from the table above that $\lor E$ requires three inputs; in this example it has received only one, so its application cannot be correct. A correct application of $\lor E$ requires a disjunction as input as well as two conditionals, each of which has one of the disjuncts from the disjunction as antecedent and both of which have the same consequent.

**Problems with $\rightarrow E$**

There are three common mistakes with $\rightarrow E$. The first is to use $\rightarrow E$ on analogy with $\&E$, using it simply to discard one side of the conditional and conclude the other, like this:

1. $P \rightarrow Q$ A
2. Show $Q$
3. $Q \rightarrow E, 1$

or like this:

1. $P \rightarrow Q$ A
2. Show $P$
3. $P \rightarrow E, 1$

Both of these are wrong. $\rightarrow E$ requires two inputs -- both a conditional and the antecedent of the conditional. With only the conditional, nothing can be inferred. Checking that you have the right number of inputs for your proof rules is an easy way to avoid this error.

The second common error with $\rightarrow E$ is to use it in reverse -- to combine a conditional with the consequent of that conditional to conclude the antecedent, like this:

1. $P \rightarrow Q$ A
2. $Q$ A
3. Show $P$
4. $P \rightarrow E, 1,2$
Again, this is wrong. A conditional is a one-way inference ticket; you can only follow it in the direction of the conditional arrow. A conditional and the consequent of that conditional do not license any conclusions using →E.

The third mistake using →E is illegitimately to use a conditional on a Show line as input to →E. This error can be tempting when you are working on some fairly complex result involving a conditional proof, like this:

- \( S \rightarrow (R & P) \)
- \( Q \rightarrow (\neg R & T) \)
- \( (Q & S) \rightarrow U \)

If you're having trouble figuring out how the proof should go, it's easy to find yourself starting a conditional proof, and then illicitly combining the content of the Show line with the ACP to magically finish the conditional proof, like this:

1. \( S \rightarrow (R & P) \) A
2. \( Q \rightarrow (\neg R & T) \) A
3. Show \( (Q & S) \rightarrow U \)
4. \( Q & S \) ACP
5. \( U \rightarrow E, 3,4 \)

This use of →E is no good because the Show of line 3 has not yet been cancelled, so that line is not yet available to serve as input to proof rules. By using the sentence on the Show line before it has been cancelled, you are without justification assuming the truth of that conditional you’re out to prove, and then using the truth of that conditional to derive the truth of that conditional. No wonder, then, that it’s so easy to complete the conditional proof. A result “proved” in this way shows nothing at all, since any result (no matter how egregiously false) could be showed in the same way. A genuine conditional proof must not use the sentence on the Show line in the process of completing that very conditional proof.

Problems with →I

There’s only one mistake you can make with the rule of →I, and that’s to use it at all. There is no rule of →I in our proof system. If you want to derive a conditional, you need to introduce a Show line declaring the intention to derive that conditional, and then carry out a conditional proof (introduce the antecedent of the conditional using ACP, and then derive the consequent of the conditional).

Problems with ¬I and ¬E

Ditto with these “rules” -- they don’t exist, so don’t use them. If you’re trying to introduce or eliminate a double negation, then use the \( \neg\neg \) rule. If you’re trying to introduce or eliminate a single negation, then you need to use the strategy of indirect proof -- write a Show line declaring an intention to derive the sentence you want (with or without the single negation), and then introduce the opposite of that sentence using AIP and derive a contradiction.

That’s it for the rule-specific problems. I’ve given no entries for &I, &E, ∨I, ↔I, and →I (as well as R, A, ACP, and AIP) because these rules are simple enough to use that I’ve never come across any student problems with them. If you do run into some difficulty with any of these rules, please let me know -- I’m always willing to add more detail.

In addition to the problems with specific rules, beginners with the proof system also often have difficulty remembering that the proof rules can only be applied when the input and output sentences have exactly the right format. When you’re using a rule associated with a particular connective, that connective must be the main connective in the relevant input or output sentence. Thus, for example, you cannot apply &E to \( \neg(P & (Q \lor R)) \), but you can apply &E to \( (P \rightarrow Q) & \neg(R \leftrightarrow (S \lor T)) \). You cannot apply ↔E to \( P \rightarrow (Q \leftrightarrow R) \), but you can apply ↔E to \( (P \rightarrow Q) \leftrightarrow (R \leftrightarrow (S \land T)) \). You cannot apply ∨I to derive \( (P \lor Q) \), but you can apply ∨I to derive \( (P \rightarrow (Q \leftrightarrow (R \land T \land U)))) \lor T \).

Here’s a quick quiz to help you catch errors. In the following proof, some of the lines (after the initial assumption lines) contain erroneous applications of the proof rules, ranging from the flagrantly obvious to the rather subtle. Try to figure out which lines have mistakes, and then check the answers.

1. P & Q A
2. P → Q A
3. Q → R A
4. \( \neg(P & Q) \leftrightarrow (Q \rightarrow P) \) A
5. Q v P A
6. \( \neg\neg Q \) A
7. \( \neg(P \lor (Q \rightarrow S)) \) A
8. (P & Q) → R A
9. Q & E, 1
10. P & E, 6
The other major source of errors in proofs is misusing the proof strategies. Misuse of proof strategy comes in three major forms:

- Invoking a proof strategy that is not permissible for the sentence being shown.
- Taking an inappropriate auxiliary assumption at the beginning of a subproof.
- Pursuing the wrong success conditions for a particular proof strategy.

First, when you write a Show line and select a proof strategy, you must be sure that the proof strategy you select is one that is permitted for the kind of claim you're trying to show. Now that we've moved to the alternative form of indirect proof, both direct and indirect proof can be used on any kind of claim. Thus the only proof strategy which has any constraints on when it can be used is conditional proof. Conditional proof can be used only when the sentence to be shown is a conditional. This means that it must have a conditional as its main connective. \(P \rightarrow Q\), \((P \& R) \rightarrow \neg (R \lor T)\), and \((P \leftrightarrow (Q \& R) \lor T) \rightarrow S\) can all be shown using conditional proof. However, \((P \rightarrow Q)\), \((P \rightarrow Q) \lor R\), and \((P \rightarrow (Q \leftrightarrow (R \& T))) \leftrightarrow (S \rightarrow U)\) cannot be shown using conditional proof, since none of them have a conditional as main connective.

Second, when you write a Show line and begin a subproof, you must be careful to take the correct auxiliary assumption. It can be quite tempting to take assumptions to which you aren't entitled. Here's a quick summary of what you are entitled to assume when you begin a subproof:

- **Direct proof**: No auxiliary assumptions
- **Conditional proof**: Assume the antecedent of the conditional to be shown using ACP.
- **Indirect proof**: Assume either (a) the unnegated form of what is to be shown (if it is negated) or (b) the negation of what is to be shown, in either case using AIP.

Probably the most frequent error is to take some sort of auxiliary assumptions when performing a direct proof. Keep in mind that you'll have no chance of picking the right auxiliary assumptions if you don't know which proof strategy you are using. Once again, every time you write a Show line, you must decide which proof strategy you are using to complete that Show.

Here's a quick practice exercise. For each of the following Show lines, indicate what auxiliary hypothesis you should take when invoking the indicated proof strategy. Then compare with the answers.

<table>
<thead>
<tr>
<th>Show Line</th>
<th>Proof Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Show ((P &amp; Q) \rightarrow R)</td>
<td>Conditional proof</td>
</tr>
<tr>
<td>(2) Show ((P \rightarrow Q))</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>(3) Show ((P \lor Q) \rightarrow R)</td>
<td>Direct proof</td>
</tr>
<tr>
<td>(4) Show ((P \lor Q) \rightarrow R)</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>(5) Show ((P \rightarrow Q) \lor (Q \rightarrow R))</td>
<td>Conditional proof</td>
</tr>
<tr>
<td>(6) Show ((P \lor Q) \rightarrow (R &amp; S))</td>
<td>Direct proof</td>
</tr>
</tbody>
</table>
The third and final problem with proof strategies is determining when you've done enough to complete the subproof triggered by a particular proof strategy. Here's a summary of the success conditions:

- **Direct proof**: Derive the sentence on the Show line.
- **Conditional proof**: Derive the consequent of the sentence on the Show line.
- **Indirect proof**: Derive a contradiction.

If you are aiming toward the wrong success conditions, you probably won't succeed in reaching them -- and even if you do, it will do you no good, because the wrong success conditions won't legitimately cancel the Show line.

For another practice exercise, try to determine for each of the following whether the lines following the Show fulfill the success conditions for the subproof, given the indicated proof strategy. Then compare with the answers.

<table>
<thead>
<tr>
<th>Show</th>
<th>Direct proof</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → Q &amp; R</td>
<td>Conditional proof</td>
</tr>
<tr>
<td>~Q &amp; R</td>
<td></td>
</tr>
<tr>
<td>P ↔ (Q ∨ S)</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>~P ↔ (Q ∨ S)</td>
<td></td>
</tr>
<tr>
<td>P ↔ (Q → (R → S))</td>
<td>Conditional proof</td>
</tr>
<tr>
<td>S</td>
<td></td>
</tr>
<tr>
<td>P ↔ (Q &amp; ~S)</td>
<td>Direct proof</td>
</tr>
<tr>
<td>~P ↔ (Q &amp; ~S)</td>
<td></td>
</tr>
<tr>
<td>~P → Q</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>~Q</td>
<td></td>
</tr>
<tr>
<td>P &amp; R</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>P → R</td>
<td></td>
</tr>
<tr>
<td>P → ~R</td>
<td></td>
</tr>
<tr>
<td>P → (R ∨ S)</td>
<td>Direct proof</td>
</tr>
<tr>
<td>R ∨ S</td>
<td></td>
</tr>
<tr>
<td>~P</td>
<td></td>
</tr>
<tr>
<td>Q ↔ S</td>
<td></td>
</tr>
<tr>
<td>~P ↔ S</td>
<td></td>
</tr>
<tr>
<td>((P ↔ Q) → S) → T</td>
<td>Conditional proof</td>
</tr>
<tr>
<td>T</td>
<td></td>
</tr>
<tr>
<td>~P</td>
<td>Indirect proof</td>
</tr>
<tr>
<td>~P</td>
<td></td>
</tr>
</tbody>
</table>

**Strategies for Proofs**

The first step in mastering the proof system is eliminating errors in the use of the proof rules. But, of course, just because you can move the pieces correctly, you're not a grandmaster at chess. You need strategy, in addition to simple competence with the rules, to master chess. Similarly, you need strategy to be good at constructing proofs. The most frequent problem beginners have with the proof system is that they simply reach a point where they have no idea what to do next in order to move the proof forward. (It's at this point that the temptation to abuse one of the rules to make it do what you need done becomes great). So in this section, we are going to go over some of the most helpful strategies for constructing proofs, to help you develop a bag of tools that will see you through the vast majority of proofs.
The first thing to note is that there are two structural features of proofs, attention to which will do about 90% of the work of guiding you through proofs. The first structural feature is that most proofs can be thought of as a nested hierarchy of conditional proofs, indirect proofs, and dilemma reasonings (with the occasional biconditional introduction thrown in). Consider, for example, the proof of:

\[ \begin{align*} &P \rightarrow S, \ R \rightarrow (Q \rightarrow S), \ \therefore (P \lor Q) \rightarrow (R \rightarrow S) \\
\end{align*} \]

The proof proceeds as follows:

\[
\begin{array}{l|c}
(1) & P \rightarrow \neg R & A \\
(2) & R \rightarrow (Q \rightarrow S) & A \\
(3) & \text{Show } (P \lor Q) \rightarrow (R \rightarrow S) & \\
(4) & P \lor Q & ACP \\
(5) & \text{Show } P \rightarrow (R \rightarrow S) & \\
(6) & P & ACP \\
(7) & \text{Show } R \rightarrow S & \\
(8) & R & ACP \\
(9) & \text{Show } S & \\
(10) & \neg S & ACP \\
(11) & R, \ S & \\
(12) & \rightarrow E, 1,8 \\
(13) & \text{Show } Q \rightarrow (R \rightarrow S) & \\
(14) & Q & ACP \\
(15) & \text{Show } R \rightarrow S & \\
(16) & R & ACP \\
(17) & Q \rightarrow S & \rightarrow E, 2,16 \\
(18) & S & \rightarrow E, 17,14 \\
(19) & R \rightarrow S & \lor E, 4,5,13 \\
\end{array}
\]

We can give a structural map of this proof as follows:

<table>
<thead>
<tr>
<th>Conditional Proof</th>
<th>Dilemma Reasoning</th>
<th>Conditional Proof</th>
<th>Indirect Proof</th>
<th>Conditional Proof</th>
<th>Conditional Proof</th>
</tr>
</thead>
</table>

In this 19-line proof, there is only one line (line 17) that is not an immediate part of this structural map. Every other line in the proof is either (a) one of the initial assumptions, (b) a Show line triggering one of the subproofs, (c) a line giving an auxiliary assumption for a subproof, or (d) a line deriving the success condition for some subproof. If you have enough skill with the particular techniques of performing conditional, indirect, and dilemma reasoning proofs, then all you have to do is to chain these techniques together appropriately, and almost the entire proof falls into place immediately.

Quite complicated proofs can actually be very simple once you see how they are hierarchically constructed out of multiple uses of a small catalogue of techniques (conditional proof, indirect proof, dilemma reasoning). It's one thing to look at the 19 lines above in toto and try to figure out what's going on. It's another (much easier) thing to see a simple conditional proof running from lines 15 to 18, and another simple conditional proof running from lines 13 to 18 (with lines 15 to 18 already accounted for), and a simple indirect proof running from lines 9 to 12, and so on. Each of the individual stages in the hierarchical structure of the proof is really quite simple, and usually contains only a few lines genuinely of its own (not counting lines of subproofs spawned by that level in the hierarchy). So if you can see the proof in two stages -- first as a structural map of subproof techniques, and second as the details of implementing each of these subproofs -- and if you can feel perfectly familiar and comfortable with the relatively minimal details of a particular strategy implementation, the proof can be built up out of small and simple pieces quite easily. This is one point at which the modular structure of the proof system is particularly advantageous.

As a practice exercise, try constructing each of the following proofs using the accompanying structural map as a guide. You should be able to put the proofs together quite easily by filling in each line in the structural map with an appropriate instance of the proof strategy in question.

**First Exercise:**

- \[ P \rightarrow (Q \rightarrow R), \ \therefore Q \rightarrow (\neg R \rightarrow \neg P) \]

**Structural Map:**
Second Exercise:
- \( S \rightarrow \neg Q, (P \lor R) \rightarrow \neg S, S, \therefore \neg (P \lor Q) \)

Structural Map:
- Indirect Proof
- Dilemma Reasoning
- Conditional Proof
- Conditional Proof
- Indirect Proof

Third Exercise:
- \((P \& Q) \lor (Q \& R), \therefore \neg Q \rightarrow S\)

Structural Map:
- Conditional Proof
- Dilemma Reasoning
- Conditional Proof
- Conditional Proof
- Indirect Proof

Of course, it's one thing to use your mastery of the particular proof strategies to understand or reconstruct a proof when you're given a structural map of it ahead of time. It's quite another to be left at the beginning of the proof and have to build the structural map yourself. The second major technique in constructing proofs, then, is being able to anticipate the direction in which the proof will typically proceed.

The most useful technique for anticipating the direction of a proof is to keep track of:
- What you know so far
and:
- What you need next
at each point in the proof, and then to have a collection of techniques for:
- Using what you know so far
and:
- Getting what you need next.

Let's look at a very simple example of this to help get the idea. Suppose we want to prove the following result:
- \( P \& T, (P \& Q) \rightarrow R, (P \lor S) \rightarrow Q, R \rightarrow S, \therefore S \lor V \)

At the beginning of the proof, we know the following:
- \( P \& T \)
- \((P \& Q) \rightarrow R\)
- \((P \lor S) \rightarrow Q\)
- \( R \rightarrow S \)
and we want to get:
- \( S \lor V \)

We can thus pursue two different lines of thought. First, we can think about how we can make forward progress from what we already know. Forward progress is typically made by using elimination rules on what we already know. Three of the sentences we already know are conditionals, and would thus call for using \( \rightarrow E \). However, using \( \rightarrow E \) would require also having the antecedent of the conditional, and since we don't have any of the antecedents in question, we can't yet make forward progress. We can, however, add all of their antecedents to the list of things we'd like to get. One of the sentences we already know is a conjunction, and would thus call for \&E. This, of course, we can do. We thus get:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>P &amp; T</td>
</tr>
<tr>
<td>(2)</td>
<td>(P &amp; Q) \rightarrow R</td>
</tr>
<tr>
<td>(3)</td>
<td>(P \lor S) \rightarrow Q</td>
</tr>
<tr>
<td>(4)</td>
<td>R \rightarrow S</td>
</tr>
<tr>
<td>(5)</td>
<td>Show S \lor V</td>
</tr>
<tr>
<td>(6)</td>
<td>P &amp;E, 1</td>
</tr>
</tbody>
</table>
What we want:

P & Q
P ∨ S
R
S ∨ V

Having thought about what we already had and how we could use it, and having thereby added two items to the list of things we know and three items to the list of things we’d like to know, we can now do one of two things:

(a) Look at the new things we know, and see if we can make any forward progress from them.
(b) Look at the things we’d like to know, and see if we can make any backward progress from them.

Both of these are worth trying, and there’s no real rule about which order they should be attempted in.

Somewhat randomly, let’s pick up option (b) and think about how we could get to any of the things we want. We are in essence trying to construct the proof backward from what we want to know (if things go well, our forward construction and our backward construction will meet somewhere in the middle). Backward progress is typically made by figuring out how to use introduction rules to get what we want to know. Here’s a quick rundown of how introduction rules could be profitably used:

- To get P & Q: Use &I, with inputs of P and Q.
- To get P ∨ S: Use ∨I, with an input of P or an input of S.
- To get R: Introduction rules give us no way of getting a single sentence letter.
- To get S ∨ V: Use ∨I, with an input of S or an input of V.

We can now use this information to build in another layer at the end of our proof:

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>(1) P &amp; T</td>
<td>(2) (P &amp; Q) → R</td>
<td>(3) (P ∨ S) → Q</td>
<td>(4) R → S</td>
<td>(5) Show S ∨ V</td>
<td>(6) P</td>
<td>(7) T</td>
</tr>
</tbody>
</table>

What we want:

P
Q
P or S, either one
S or V, either one
P & Q
P ∨ S
R
S ∨ V

The list of what we need (and don’t already have a plan for getting via introduction rules) is now:

- P
- Q
- R
- S
- V

There’s some redundancy in this list (we don’t need both P and S, for example), but we would need some of these to move on to the next phase of what we need. That next phase would either give us our final conclusion S ∨ V (if we managed to get S or V) or would give us (or contribute to giving us) the antecedent of one of our conditionals (if we got P, Q, R, or S).

The updated list of what we need consists entirely of sentence letters, and single sentence letters can’t be obtained using introduction rules. So at this point it’s time to stop working backward (thinking about what we need). We could now go back to the list of what we already have, and see if elimination rules could be applied to any of those things to make more forward progress. However, it turns out that that we can do something even more useful, because at this point the list of what we want and the list of what we already have overlap. We already have P, from the application of &E to P & T. And
we want $P$, for two purposes -- first, to combine with $Q$ via $\&I$ to make $P \& Q$, and second, to apply $\lor I$ to make $P \lor S$.

We can therefore go ahead and apply $\lor I$ to achieve the second of these purposes:

| (1) $P \& T$ | A |
| (2) $(P \& Q) \rightarrow R$ | A |
| (3) $(P \lor S) \rightarrow Q$ | A |
| (4) $R \rightarrow S$ | A |
| (5) Show $S \lor V$ |  |
| (6) $P$ | $\&E, 1$ |
| (7) $T$ | $\&E, 1$ |
| (8) $P \lor S$ | $\lor I, 6$ |

...What we want:

| $Q$ |  |
| $S \lor V$, either one |  |
| $P \& Q$ | $\&I$ |
| $R$ |  |
| $S \lor V$ | $\lor I$ |

I've now removed both $P$ and either-$P$-or-$S$ from the list of things we want (but left $Q$, which at some point we'd like to combine with $P$ to make $P \& Q$), as well as $P \lor S$.

Now that we've added to our list of things that we know, it's a good idea to look back over that list and see if we can make further forward progress. Here's the list of what we know so far:

- $P \& T$
- $(P \& Q) \rightarrow R$
- $(P \lor S) \rightarrow Q$
- $R \rightarrow S$
- $P$
- $T$
- $P \lor S$

Again, the thought is that we will use elimination rules on what we already know to make forward progress. Since there is no elimination rule for a single sentence letter, this leaves us with three kinds of options:

- Apply $\rightarrow E$ to $(P \& Q) \rightarrow R$, $(P \lor S) \rightarrow Q$, or $R \rightarrow S$
- Apply $\&E$ to $P \& T$
- Apply $\lor E$ to $P \lor S$

We've already pursued the second option, of course, so it's not going to produce anything new for us. The third option would require us to use $\lor E$ and do dilemma reasoning, which is a pain, so let's hold off on that until we have no other choice (one good piece of advice for constructing proofs: always take the easiest available course). So we'll look at the second option (that of using $\rightarrow E$ on one of our conditionals). Remember that to use $\rightarrow E$ we need the antecedent of one of the conditionals. However, we've just added $P \lor S$ to the list of things we know, and $P \lor S$ is the antecedent of $(P \lor S) \rightarrow Q$. Thus we're all set to move forward with $\rightarrow E$.

Here's a quicker way of reaching the same point: We just removed $P \lor S$ from the list of things that we wanted because we had succeeded in deriving it from $P$ using $\lor I$. But, of course, $P \lor S$ was on the list of things we wanted for a reason -- we put it there earlier because we saw that we could combine it with $(P \lor S) \rightarrow Q$ to make forward progress. So once we get it, we know we're in a position to put it to work. The general moral: any time you obtain something from the list of things you want, remember why you've got it on that list, and then go ahead and use it for that purpose.

Applying $\rightarrow E$ to $P \lor S$ and $(P \lor S) \rightarrow Q$, we get:

| (1) $P \& T$ | A |
| (2) $(P \& Q) \rightarrow R$ | A |
| (3) $(P \lor S) \rightarrow Q$ | A |
| (4) $R \rightarrow S$ | A |
| (5) Show $S \lor V$ |  |
| (6) $P$ | $\&E, 1$ |
| (7) $T$ | $\&E, 1$ |
| (8) $P \lor S$ | $\lor I, 6$ |
What we want:
S or V, either one
P & Q & I
R
S ∨ V

I've removed Q from the list of things we want, since we now have it. We wanted Q so that we could combine it with P to make P & Q, which is also on the list of things we want. We're now in a position to do that, so we have:

(1) P & T
(2) (P & Q) → R
(3) (P ∨ S) → Q
(4) R → S
(5) Show S ∨ V
(6) P &E, 1
(7) T &E, 1
(8) P ∨ S ∨ I, 6
(9) Q →E, 3,8
(10) P & Q &I, 6,9
...
What we want:
S or V, either one
R
S ∨ V

We've now removed P & Q from the list of things that we want, so we go back and check why we wanted it. We find that P & Q was wanted to let us make forward progress from (P & Q) → R using →E, so we now do that:

(1) P & T
(2) (P & Q) → R
(3) (P ∨ S) → Q
(4) R → S
(5) Show S ∨ V
(6) P &E, 1
(7) T &E, 1
(8) P ∨ S ∨ I, 6
(9) Q →E, 3,8
(10) P & Q &I, 6,9
(11) R →E, 2,10
...
What we want:
S or V, either one
S ∨ V

Now we've removed R from the list of things we want, so we go back and see that we wanted R to allow us to make forward progress from R → S using →E. Doing that, we obtain:

(1) P & T
(2) (P & Q) → R
(3) (P ∨ S) → Q
(4) R → S
(5) Show S ∨ V
(6) P &E, 1
(7) T &E, 1
(8) P ∨ S ∨ I, 6
(9) Q →E, 3,8
(10) P & Q &I, 6,9
(1) P & T \\
(2) (P & Q) → R \\
(3) (P ∨ S) → Q \\
(4) R → S \\
(5) Show S ∨ V \\
(6) P &E, 1 \\
(7) T &E, 1 \\
(8) P ∨ S ∨I, 6 \\
(9) Q →E, 3,8 \\
(10) P & Q &I, 6,9 \\
(11) R →E, 2,10 \\
(12) S →E, 4,11 \\
(13) S ∨ V ∨I, 12

The important point to note here is that the entirety of this proof falls into place mechanically just by alternatingly working forward from what we already know and working backward from what we want to get. Thus we get a very general strategy for approaching proofs:

Try to work forward by applying elimination rules to what you already know, and try to work backward by figuring out how to use introduction rules to get what you want. Work from both ends until they meet in the middle, and then paste the two halves together.

Notice how the modular structure of the proof rules supports this strategy. Broadly speaking, for each connective we have both an introduction rule and an elimination rule, so we have one way to move forward from a sentence of a certain type (i.e., with a certain main connective) and one way to work backward from a sentence of a certain type.

It's worth noting that this back-and-forth method won't always produce the most efficient proof (in fact, as we'll see below it won't always produce a proof at all), because sometimes some of the paths you'll follow in either the forward or the backward direction turn out to be superfluous. For example, in the above proof we derive T from P & T as part of the forward movement from P & T, but T is never in fact used in the proof – the proof would be just as valid if we omitted line 7. As you get better at proofs, it will be easier to see ahead of time which paths are worth exploring, but it's a good idea at first just to get used to the idea of tracing out every forward and backward line, so that you'll be able to find the places where the lines meet.

An Example of Applying Proof Strategies [Next]

Let's look at a more complicated example now, to see how the same method applies. Recall the earlier proof of:

- P → S, R → (Q → S), ∴ (P ∨ Q) → (R → S)

Suppose that instead of being given the structural map of this proof (as you were before), you were left on your own to figure it out. Let's see how we can combine the work-from-both-ends forward-and-backward method with the modular structural map method to tackle a complex proof.

The proof, of course, begins like this:

(1) P → ¬R \\
(2) R → (Q → S) \\
(3) Show (P ∨ Q) → (R → S)

At this point, there are two things on our list of what we know:

- P → ¬R
- R → (Q → S)

and one thing on our list of what we'd like to have:

- (P ∨ Q) → (R → S)
Since both of the things we know are conditionals, moving forward involves the →E rule, and hence requires us to know the antecedents of the conditionals. We don’t yet know either of these, so we’ll add them to the list of things we’d like, to get this:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P → ¬R A</td>
</tr>
<tr>
<td>2</td>
<td>R → (Q → S) A</td>
</tr>
<tr>
<td>3</td>
<td>Show (P ∨ Q) → (R → S)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What we want:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td></td>
</tr>
<tr>
<td>(P ∨ Q) → (R → S)</td>
<td></td>
</tr>
</tbody>
</table>

Our other option is to work backward from what we want. We can’t work backward from P or R, since each of these is a simple sentence letter. However, we can try to work backward from (P ∨ Q) → (R → S). Since this sentence is a conditional, and since we work backward using introduction rules, it seems that we ought to appeal to conditional introduction here. However, there is no rule of conditional introduction.

Instead, we invoke conditional proof. Whenever we want to get some sentence of a form for which there is no introduction rule, we instead invoke the appropriate proof strategy. Thus when we want a conditional, we invoke conditional proof. When we want a negation, we invoke indirect proof.

Invoking conditional proof does two things. First, it adds one item to what we know – namely, the antecedent of the conditional, taken by ACP. Second, it adds one item to what we want – namely, the consequent of the conditional. It also places both this new information and this new desire within the conditional subproof, to maintain proper informational hygiene. We thus have:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P → ¬R A</td>
</tr>
<tr>
<td>2</td>
<td>R → (Q → S) A</td>
</tr>
<tr>
<td>3</td>
<td>Show (P ∨ Q) → (R → S)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>P ∨ Q ACP</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What we want:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R → S</td>
<td></td>
</tr>
</tbody>
</table>

Finding a conditional among the things we wanted to know thus triggered the first entry on the structural map of the proof – we know that there is a conditional proof in the larger proof, and we thus invoke our mastery of the conditional proof form to fill in the outlines of what we know must happen.

Now we move on. We’ve added one new item to the list of what we know, so that list now includes:

- P → ¬R
- R → (Q → S)
- P ∨ Q

We’ve also added one item to the list of what we want to have, so that list now includes:

- R → S
- P
- R
- [(P ∨ Q) → (R → S)]

I’ve bracketed (P ∨ Q) → (R → S), because we’ve already worked out how to get that – as soon as we get R → S, which is also on the wish list, then we automatically get (P ∨ Q) → (R → S) as well. Thus there’s no reason to worry any more about (P ∨ Q) → (R → S).

As usual, we’ve got two choices at this point. We can either try to move forward from what we already know, or move backward from what we want to know. There’s no general ground for choosing between these two options. Sometimes only one of the two will actually yield any progress, and then it’s easy. Sometimes both are live options, and then you just
have to pick one. Frequently one choice will lead to a shorter final proof than the other (although it's quite difficult to see ahead of time which), but both will get you there in the end. I'll arbitrarily choose to try moving forward, rather than working backward.

We've already tried moving forward from our two conditionals, and found that we can't do so until we find the antecedents of those conditionals (which we haven't done yet). All that's left, then, is the disjunction $P \lor Q$. Moving forward from a disjunction requires, of course, using the elimination rule $\lor E$. Unfortunately, the strategy of using $\lor E$ is a bit more complicated than the other rules we've used so far. The output of $\lor E$ is wholly unconstrained by the starting disjunction, since the rule takes inputs of form $\Phi \lor \Theta$, $\Phi \to \Psi$, and $\Theta \to \Psi$ and produces an output of $\Psi$. We know, by way of having $P \lor Q$, what $\Phi$ and $\Theta$ are, but $\Psi$ can be anything we want. Of course, it makes sense to think that we should pick for $\Psi$ one of the three sentences we want, but which one? There's no easy answer to that question; frequently it's just a matter of trying until something works.

I'm going to let $\Psi$ be $R \to S$, both because (a) I happen to know already that that will work, and (b) when in doubt, you may as well go for the big game, and $R \to S$ would give us what we needed to bring the proof to a close, whereas both $P$ and $R$ would only allow us to use $\to E$ on a premise. Now that we've picked $\Psi$, $\lor E$ (in its usual helpful way) does not in fact add anything to the list of what we know, but instead adds two things to the list of what we want – the two conditionals $\Phi \to \Psi$ and $\Theta \to \Psi$. Our proof thus looks like thus so far:

| (1) $P \to \neg R$ | A |
| (2) $R \to (Q \to S)$ | A |
| (3) Show $(P \lor Q) \to (R \to S)$ |
| (4) $P \lor Q$ | ACP |
| ... |
| What we want: |
| $P \to (R \to S)$ |
| $Q \to (R \to S)$ |
| $R \to S$ |
| ... |
| What we want: |
| $P$ |
| $R$ |
| $(P \lor Q) \to (R \to S)$ |

where the two lines in blue have been added by the appeal to $\lor E$. Notice that the additions are, like $P \lor Q$, within the conditional subproof. We've now triggered the second entry (a dilemma reasoning) on the structural map of the proof, which now looks like this:

| Dilemma Reasoning |
| Conditional Proof |

Since we haven't added anything new to our list of known information, working forward has for the moment come to a dead end. It's time, then, to try working backward. Here's the current list of what we want:

- $P \to (R \to S)$
- $Q \to (R \to S)$
- $[R \to S]$
- $P$
- $R$
- $[(P \lor Q) \to (R \to S)]$

As before, I've bracketed sentences for which we already have a plan in place. This leaves $P$ and $R$, which we still can't get using introduction rules, and the two conditionals $P \to (R \to S)$ and $Q \to (R \to S)$. We are forced, then, to try to get one of those conditionals. Let's pick – entirely at random – $P \to (R \to S)$. Again, we want to use an introduction rule when working backward, but there is no rule of conditional introduction, so we invoke conditional proof instead. Doing so gives us this:

| (1) $P \to \neg R$ | A |
| (2) $R \to (Q \to S)$ | A |
| (3) Show $(P \lor Q) \to (R \to S)$ |
| (4) $P \lor Q$ | ACP |
| (5) Show $P \to (R \to S)$ |
| (6) $P$ | ACP |
What we want:
- \( R \rightarrow S \)

What we want:
- \( P \rightarrow (R \rightarrow S) \)
- \( Q \rightarrow (R \rightarrow S) \)
- \( R \rightarrow S \)

What we want:
- \( P \)
- \( R \)
- \( (P \lor Q) \rightarrow (R \rightarrow S) \)

Making this move has added \( P \) to the list of things we know, and added \( R \rightarrow S \) to the list of things we want (more properly, "reactivated" \( R \rightarrow S \)) in that list. Making this move also adds a conditional proof to the structural map, giving us:

<table>
<thead>
<tr>
<th>Conditional Proof</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilemma Reasoning</td>
</tr>
<tr>
<td>Conditional Proof</td>
</tr>
</tbody>
</table>

At this point, we have an overlap between the list of things we know and the list of things we want – \( P \) occurs on both. We thus remind ourselves why we wanted \( P \) (to use \( \rightarrow E \) on \( P \rightarrow \neg R \)), and act accordingly, thus yielding:

<table>
<thead>
<tr>
<th>(1) ( P \rightarrow \neg R )</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) ( R \rightarrow (Q \rightarrow S) )</td>
<td>A</td>
</tr>
<tr>
<td>(3) Show ( (P \lor Q) \rightarrow (R \rightarrow S) )</td>
<td>ACP</td>
</tr>
<tr>
<td>(4) ( P \lor Q )</td>
<td>ACP</td>
</tr>
<tr>
<td>(5) Show ( P \rightarrow (R \rightarrow S) )</td>
<td>ACP</td>
</tr>
<tr>
<td>(6) ( P )</td>
<td>ACP</td>
</tr>
<tr>
<td>(7) ( \neg R )</td>
<td>( \rightarrow E, 1,6 )</td>
</tr>
</tbody>
</table>

What we want:
- \( R \rightarrow S \)

... What we want:
- \( P \rightarrow (R \rightarrow S) \)
- \( Q \rightarrow (R \rightarrow S) \)
- \( R \rightarrow S \)

Things we know:
- \( P \rightarrow \neg R \)
- \( R \rightarrow S \)
- \( R \rightarrow (Q \rightarrow S) \)
- \( [P \rightarrow (R \rightarrow S)] \)
- \( P \lor Q \)
- \( Q \rightarrow (R \rightarrow S) \)
- \( P \)
- \( R \)
- \( \neg R \)
- \( [(P \lor Q) \rightarrow (R \rightarrow S)] \)

The only new item on the list of things we know is \( \neg R \), and there is unfortunately no elimination rule for negation (the closest we have is \( \neg \neg \), but that is obviously not applicable in the case of \( \neg R \)). It's time, then, to go back to the list of things we want to know and work backward. The two obvious choices are \( R \rightarrow S \) and \( Q \rightarrow (R \rightarrow S) \). Either one will work (as long as we situate it correctly in the subproof structure), but I'll choose \( R \rightarrow S \) for now, so that we can continue to pursue the subproof of \( P \rightarrow (R \rightarrow S) \). We thus add a conditional proof outline to get:

| (1) \( P \rightarrow \neg R \) | A |
We've now added \( R \) to the list of things we know and \( S \) to the list of things we want. Since both of these are single sentence letters, we don't have any introduction or elimination rules for either. However, our newly-discovered \( R \) does coincide with one of the things we wanted, and puts us in position to use \( \rightarrow E \) on our second assumption, to get:
This adds \( Q \to S \) to the list of things we know, and also prompts us to add \( Q \) (the antecedent of \( Q \to S \), and thus what we need to move forward from \( Q \to S \) using \( \to \text{E} \)) to the list of things we want. Our updated lists are:

<table>
<thead>
<tr>
<th>Things we know:</th>
<th>Things we want:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P \to \neg R )</td>
<td>( R \to S )</td>
</tr>
<tr>
<td>( R \to (Q \to S) )</td>
<td>( [P \to (R \to S)] )</td>
</tr>
<tr>
<td>( P \lor Q )</td>
<td>( Q \to (R \to S) )</td>
</tr>
<tr>
<td>( P )</td>
<td>( R )</td>
</tr>
<tr>
<td>( \neg R )</td>
<td>( [(P \lor Q) \to (R \to S)] )</td>
</tr>
<tr>
<td>( Q \to S )</td>
<td>( Q )</td>
</tr>
<tr>
<td>( R )</td>
<td>( S )</td>
</tr>
</tbody>
</table>

**Question:** Why didn't you remove \( R \) from the list of things we want, since we've now gotten and used it?

**Answer:** We did get \( R \), but its presence and application is limited to the subproof of \( R \to S \). Since \( R \to (Q \to S) \), the conditional that prompted us to want \( R \), is not limited to that subproof, it's possible that we'll want \( R \) again later, to use that conditional again after we complete this subproof.

We're stuck again in the forward direction, so we turn back to things we want. We've just added both \( Q \) and \( S \) to this list, but we can't pursue either of those with introduction rules. Thus we have to look back a few steps, and take up the path of trying to get \( Q \to (R \to S) \). Here we'll use conditional proof again, yielding:

\[
\begin{align*}
(1) & \quad P \to \neg R \\
(2) & \quad R \to (Q \to S) \\
(3) & \quad \text{Show } (P \lor Q) \to (R \to S) \\
(4) & \quad P \lor Q \quad \text{ACP} \\
(5) & \quad \text{Show } P \to (R \to S) \\
(6) & \quad P \quad \text{ACP} \\
(7) & \quad \neg R \quad \to \text{E}, 1,6 \\
(8) & \quad \text{Show } R \to S \\
(9) & \quad R \quad \text{ACP} \\
(10) & \quad Q \to S \quad \to \text{E}, 2,9 \\
\ldots & \quad \text{What we want:} \\
& \quad S \\
\ldots & \quad \text{What we want:} \\
& \quad R \to S \\
\end{align*}
\]

\[
\begin{align*}
(1) & \quad \text{Show } Q \to (R \to S) \\
(1) & \quad Q \quad \text{ACP} \\
\ldots & \quad \text{What we want:} \\
& \quad R \to S \\
\ldots & \quad \text{What we want:} \\
& \quad P \to (R \to S) \\
& \quad Q \to (R \to S) \\
& \quad R \to S \\
\ldots & \quad \text{What we want:} \\
& \quad R \\
(1) & \quad (P \lor Q) \to (R \to S) \\
\end{align*}
\]

and a structural map of:

**Conditional Proof**
Question: Why do you have empty parenthesis () with no line numbers before Show Q → (R → S) and Q?

Answer: Because until we finish the subproof of P → (R → S), we don't know how long it will be, and hence on which line the subproof of Q → (R → S) will begin. Once we do finish that earlier subproof, we'll go back and fill in the missing line numbers.

We've now given ourselves Q to work with, and R → S as a target. Q is actually on our list of things we want, but it's wanted for the conditional Q → S in the subproof of P → (R → S), and thus the instance of Q we just got can't be used to satisfy that want. (This is an illustration of how important it is to keep the structural map of the proof, as well as the forward-and-backward list, in mind.) However, the target of R → S we can use, thereby triggering another conditional proof:

| 1) P → ¬R                        | A                      |
| 2) R → (Q → S)                  | A                      |
| 3) Show P ∨ Q → (R → S)         |                         |
| 4) P ∨ Q                       | ACP                    |
| 5) Show P → (R → S)             |                         |
| 6) P                           | ACP                    |
| 7) ¬R                          | →E, 1, 6               |
| 8) Show R → S                  |                         |
| 9) R                           | ACP                    |
| 10) Q → S                      | →E, 2, 9               |
| ...                            |                         |
| What we want:                  |                         |
| S                              |                         |
| ...                            |                         |
| What we want:                  |                         |
| R → S                          |                         |
| () Show Q → (R → S)             |                         |
| () Show R → S                  |                         |
| R                              | ACP                    |
| ...                            |                         |
| What we want:                  |                         |
| S                              |                         |
| ...                            |                         |
| What we want:                  |                         |
| R → S                          |                         |
| ...                            |                         |
| What we want:                  |                         |
| P → (R → S)                    |                         |
| Q → (R → S)                    |                         |
| R → S                          |                         |
| ...                            |                         |
| What we want:                  |                         |
| R                              |                         |
| (P ∨ Q) → (R → S)              |                         |

and a new structural map:

Conditional Proof
We've now added $R$ to our list of what we've got, and $S$ to our wish list. As in the subproof of $P \rightarrow (R \rightarrow S)$, getting $R$ allows us to move forward from $R \rightarrow (Q \rightarrow S)$ using $\rightarrow E$, to get:

|   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | $P \rightarrow \neg R$ | A |
| (2) | $R \rightarrow (Q \rightarrow S)$ | A |
| (3) Show $(P \lor Q) \rightarrow (R \rightarrow S)$ |   |
| (4) | $P \lor Q$ | ACP |
| (5) Show $P \rightarrow (R \rightarrow S)$ |   |
| (6) | $P$ | ACP |
| (7) | $\neg R$ | $\rightarrow E$, 1, 6 |
| (8) Show $R \rightarrow S$ |   |
| (9) | $R$ | ACP |
| (10) | $Q \rightarrow S$ | $\rightarrow E$, 2, 9 |

... What we want:

$R \rightarrow S$

(1) Show $Q \rightarrow (R \rightarrow S)$

| (1) | $Q$ | ACP |
| (2) Show $R \rightarrow S$ |   |
| (3) | $R$ | ACP |
| (4) | $Q \rightarrow S$ | $\rightarrow E$, 2, 1 |

... What we want:

$S$

... What we want:

$R \rightarrow S$

... What we want:

$P \rightarrow (R \rightarrow S)$

$Q \rightarrow (R \rightarrow S)$

$R \rightarrow S$

... What we want:

$R$

$(P \lor Q) \rightarrow (R \rightarrow S)$

As when we derived $Q \rightarrow S$ within the subproof of $P \rightarrow (R \rightarrow S)$, we now add both $Q$ and $S$ to our wish list. However, within this subproof, we have $Q$ on the list of things we know, so we can move forward from $Q \rightarrow S$ using $\rightarrow E$. Doing so gives us $S$, so we can remove $S$ from the list of things we want. But that's not all – getting $S$ completes the conditional proof of $R \rightarrow S$, so we can remove it from our wish list. And getting $R \rightarrow S$ completes our conditional proof of $Q \rightarrow (R \rightarrow S)$, so we can remove it from our wish list. Clearing things out, we get:

|   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | $P \rightarrow \neg R$ | A |
| (2) | $R \rightarrow (Q \rightarrow S)$ | A |
| (3) Show $(P \lor Q) \rightarrow (R \rightarrow S)$ |   |
| (4) | $P \lor Q$ | ACP |
| (5) Show $P \rightarrow (R \rightarrow S)$ |   |
| (6) | $P$ | ACP |
| (7) | $\neg R$ | $\rightarrow E$, 1, 6 |
Let's recap now what we know and what we want:

<table>
<thead>
<tr>
<th>Things we know:</th>
<th>Things we want:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P \rightarrow \neg R )</td>
<td>( [R \rightarrow S] )</td>
</tr>
<tr>
<td>( R \rightarrow (Q \rightarrow S) )</td>
<td>( [P \rightarrow (R \rightarrow S)] )</td>
</tr>
<tr>
<td>( P \lor Q )</td>
<td>( [P \lor Q \rightarrow (R \rightarrow S)] )</td>
</tr>
<tr>
<td>( P )</td>
<td>( Q )</td>
</tr>
<tr>
<td>( \neg R )</td>
<td>( S )</td>
</tr>
<tr>
<td>( Q \rightarrow S )</td>
<td></td>
</tr>
<tr>
<td>( R )</td>
<td></td>
</tr>
<tr>
<td>( Q \rightarrow (R \rightarrow S) )</td>
<td></td>
</tr>
</tbody>
</table>

Things have come rather to a standstill at this point. Our only active wants are \( Q \) and \( S \), and these are both simple sentence letters, and can't be gotten through any introduction rule. Nor can any of what we know be used. All of the conditionals have either been used already or have antecedents we don't have. \( P \lor Q \), of course, we are in the process of using. And \( P \), \( R \), and \( \neg R \) don't have any corresponding elimination rules.

We need a new idea. I'll thus suggest the following principle:

- **When all else fails, try an indirect proof to get something you want.**

  The two things we want are \( Q \) and \( S \), so we could try an indirect proof on either of them. Since the only reason we would want \( Q \) would be to get \( S \) via \( Q \rightarrow S \), it makes sense to cut out the middle man and go directly for \( S \). We thus start an indirect proof of \( S \), yielding:

<table>
<thead>
<tr>
<th>Proof steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) ( P \rightarrow \neg R )</td>
</tr>
<tr>
<td>(2) ( R \rightarrow (Q \rightarrow S) )</td>
</tr>
<tr>
<td>(3) ( P \lor Q \rightarrow (R \rightarrow S) )</td>
</tr>
<tr>
<td>(4) ( P \lor Q )</td>
</tr>
<tr>
<td>(5) ( P \rightarrow (R \rightarrow S) )</td>
</tr>
<tr>
<td>(6) ( P )</td>
</tr>
<tr>
<td>(7) ( R )</td>
</tr>
<tr>
<td>(8) ( R \rightarrow S )</td>
</tr>
<tr>
<td>(9) ( R )</td>
</tr>
<tr>
<td>(10) ( Q \rightarrow S )</td>
</tr>
<tr>
<td>(11) ( \neg R )</td>
</tr>
<tr>
<td>(12) ( \neg S )</td>
</tr>
</tbody>
</table>
What we want:

Any contradiction

What we want:

S

What we want:

R → S

<table>
<thead>
<tr>
<th>() Show Q → (R → S)</th>
<th>ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>() Q</td>
<td>ACP</td>
</tr>
<tr>
<td>() Show R → S</td>
<td></td>
</tr>
<tr>
<td>() Q → S</td>
<td>→E, 2, ()</td>
</tr>
<tr>
<td>() S</td>
<td>→E, 1, ()</td>
</tr>
</tbody>
</table>

What we want:

P → (R → S)

R → S

What we want:

P ∨ Q → (R → S)

and a new structural map:

Conditional Proof

Dilemma Reasoning

Conditional Proof

Conditional Proof

Indirect Proof

Conditional Proof

Conditional Proof

Starting this indirect proof adds ¬S to the list of things we know, and gives us a new goal of getting a contradiction. Here’s our updated list of what we have and what we want:

<table>
<thead>
<tr>
<th>Things we know:</th>
<th>Things we want:</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → ¬R</td>
<td>[R → S]</td>
</tr>
<tr>
<td>R → (Q → S)</td>
<td>[P → (R → S)]</td>
</tr>
<tr>
<td>P ∨ Q</td>
<td>[P ∨ Q) → (R → S)]</td>
</tr>
<tr>
<td>¬R</td>
<td>Q</td>
</tr>
<tr>
<td>Q → S</td>
<td>[S]</td>
</tr>
<tr>
<td>¬S</td>
<td>Any contradiction</td>
</tr>
<tr>
<td>Q → (R → S)</td>
<td></td>
</tr>
</tbody>
</table>

Now it turns out that we’re in luck. Looking over the list of what we know, we see that we’ve already got a contradiction available – R and ¬R. Furthermore, both R and ¬R are available for downward transmission. So a couple of uses of the reiteration rule will put them in the indirect proof and finish that proof. Finishing it will give us S, and hence give us R → S, and hence give us P → (R → S). Updating our proof-in-progress, we get:

| (1) P → ¬R | A |
| (2) R → (Q → S) | A |
| (3) Show (P ∨ Q) → (R → S) | |
| (4) P ∨ Q | ACP |
| (5) Show P → (R → S) | |
| (6) P | ACP |
| (7) ¬R | →E, 1, 6 |
(Notice that I've filled in the missing line numbers, now that the two fragments of the proof have connected.) We've now added $P \rightarrow (R \rightarrow S)$ to our list of what we know. Both $P \rightarrow (R \rightarrow S)$ and $Q \rightarrow (R \rightarrow S)$ were added to the wish list in the first place to allow us to move forward from $P \lor Q$, so now that we've got both it's finally time to use $\lor E$ to get:

(1) $P \rightarrow \neg R$  
(2) $R \rightarrow (Q \rightarrow S)$  
(3) Show $P \lor Q \rightarrow (R \rightarrow S)$  
(4) $P \lor Q$  
(5) Show $P \rightarrow (R \rightarrow S)$  
(6) $P$  
(7) $\neg R$  
(8) Show $R \rightarrow S$  
(9) $R$  
(10) $Q \rightarrow S$  
(11) Show $S$  
(12) $\neg S$  
(13) $R$  
(14) $\neg R$  
(15) Show $Q \rightarrow (R \rightarrow S)$  
(16) $Q$  
(17) Show $R \rightarrow S$  
(18) $R$  
(19) $Q \rightarrow S$  
(20) $S$  
(21) $R \rightarrow S$  

(1) $P \rightarrow \neg R$  
(2) $R \rightarrow (Q \rightarrow S)$  
(3) Show $(P \lor Q) \rightarrow (R \rightarrow S)$  
(4) $P \lor Q$  
(5) Show $P \rightarrow (R \rightarrow S)$  
(6) $P$  
(7) $\neg R$  
(8) Show $R \rightarrow S$  
(9) $R$  
(10) $Q \rightarrow S$  
(11) Show $S$  
(12) $\neg S$  
(13) $R$  
(14) $\neg R$  
(15) Show $Q \rightarrow (R \rightarrow S)$  
(16) $Q$  
(17) Show $R \rightarrow S$  
(18) $R$  
(19) $Q \rightarrow S$  
(20) $S$  
(21) $R \rightarrow S$  

Notice that getting $R \rightarrow S$ from $\lor E$ also gives us what we need to finish the conditional proof of $(P \lor Q) \rightarrow (R \rightarrow S)$, which in turn brings the entire proof to a close.

Proof Strategies Summarized [Next]
OK, let's review what we've got so far. I have suggested a two-pronged strategy for constructing proofs. First, the step-by-step details of the proof will be slotted into place by pursuing a strategy of working both forward from what we know and backward from what we'd like to know. In doing this, you can reference the following chart giving advice on how to do each, depending on what material you are dealing with:

<table>
<thead>
<tr>
<th>And you want to...</th>
<th>Go forward</th>
<th>Work backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>If the main...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>¬</td>
<td>Wait for the Negated Sentence to Show Up on the List of Things You Need. Note: Moving Forward from Negations Is Often Difficult.</td>
<td>Use an Indirect Proof to Get the Negated Sentence. Therefore, Add the Unnegated Form of the Sentence to What You Know, and Any Contradiction to What You Want.</td>
</tr>
<tr>
<td>&amp;</td>
<td>Use &amp;E to Separate the Two Conjuncts. Therefore, Add Both Conjuncts to What You Know.</td>
<td>Use &amp;I on the Two Conjuncts to Derive the Conjunction. Therefore, Add Both Conjuncts to What You Want.</td>
</tr>
<tr>
<td>∨</td>
<td>Use ∨E to Move from the Disjunction to Something That You Want. Therefore, Add What You Want Two Conditionals, One From Each Conjunct to Whatever You’ve Selected from Your Wish List as the Target of the ∨E.</td>
<td>Use ∨I to Move from One of the Disjuncts to the Whole Disjunction. Therefore, Add Either of the Disjuncts to What You Want. Note: It’s Often Difficult to Derive Disjunctions via One Disjunct Using ∨I.</td>
</tr>
<tr>
<td>→</td>
<td>Use →E to Combine the Conditional with Its Antecedent to Derive the Consequent. Therefore, Add the Antecedent of the Conditional to What You Want, and When You Get That Antecedent, Add the Consequent to What You Know.</td>
<td>Use Conditional Proof to Derive the Conditional. Therefore, Add the Antecedent of the Conditional to What You Know, and Add the Consequent of the Conditional to What You Want.</td>
</tr>
<tr>
<td>↔</td>
<td>Use ↔E to Combine the Biconditional with One Side of the Biconditional to Derive the Other Side. Therefore, Add One Side of the Biconditional to What You Want, and When You Get That Side, Add the Other Side to What You Know.</td>
<td>Use ↔I to Combine Conditionals in Both Directions into a Biconditional. Therefore, Add Conditionals in Both Directions to What You Want.</td>
</tr>
</tbody>
</table>

Second, as you follow the back-and-forth method, construct a structural map of the proof and use your mastery of the techniques of conditional proof, indirect proof, and dilemma reasoning to help you construct a hierarchical sequence of those moves.

Let’s do one more example using these two strategies in conjunction, to show how much of the work can be done by them. We’ll prove the following result: 

\[ P \rightarrow R, \quad \neg R \rightarrow \neg Q, \quad \therefore (P \lor Q) \rightarrow R \]

The beginning of the proof, of course, is automatic:

\[
(1) P \rightarrow R \quad \text{A}
\]
Now we start working forward and backward. Working forward, we have two conditionals, so we want antecedents for those conditions to let us use →E. We thus add P and ¬R to our wish list. Working backward, the presence of (P ∨ Q) → R on our wish list triggers a conditional proof, like this:

What we want:
(P ∨ Q) → R

(1) P → R A
(2) ¬R → ¬Q A
(3) Show (P ∨ Q) → R
(4) P ∨ Q ACP

We've thus added P ∨ Q to what we know, and R to what we want. R doesn't help us for working backward (yet – we can always try an indirect proof of R if we get desperate), so we turn to working forward from P ∨ Q. This means using vE, which in turn means picking a target and trying to get conditionals from each disjunct to that target. The obvious target to aim for here is R, since that's what we need to finish the conditional proof. We thus add P → R and Q → R to our wish list – but immediately remove P → R, since we've already got it. We now have:

What we want:
Q → R

The only change has been the addition of Q → R to our wish list, so let's pursue that option. Q → R is a conditional, so trying to achieve it triggers a conditional proof, which adds Q to our list of what we want and R to our list of what we want. We thus have:

(1) P → R A
(2) ¬R → ¬Q A
(3) Show (P ∨ Q) → R
(4) P ∨ Q ACP
(5) Show Q → R
(6) Q ACP

What we want:
R
Our current list of what we know and what we want is:

<table>
<thead>
<tr>
<th>What we know</th>
<th>What we want</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → R</td>
<td>[(P ∨ Q) → R]</td>
</tr>
<tr>
<td>¬R → ¬Q</td>
<td>[Q → R]</td>
</tr>
<tr>
<td>P ∨ Q</td>
<td>P</td>
</tr>
<tr>
<td>Q</td>
<td>R</td>
</tr>
<tr>
<td>¬R</td>
<td>¬R</td>
</tr>
</tbody>
</table>

We're still not in a position to use P → R or ¬R → ¬Q, and we're in the process of using P ∨ Q, and there's nothing we can do right now with R. So moving forward is out. On the backward side, the active goals are P, R, and ¬R. We could choose to pursue any of these three using indirect proof, so there's a choice to be made here. I'm going to try for R, on the following ground:

* When in doubt, try to derive whatever result will be of the most help.
Since R will immediately bring the conditional proof of Q → R to an end, while neither P or ¬R is of any immediately obvious help, it's worth going for it.

Starting the indirect proof, we get:

1. P → R
2. ¬R → ¬Q
3. Show (P ∨ Q) → R
4. P ∨ Q
5. Show Q → R
6. Q
7. Show R
8. ¬R

... What we want:
Any contradiction

... What we want:
R

What we want:
Q → R
R
P
¬R
(P ∨ Q) → R

with the following structural map:

<table>
<thead>
<tr>
<th>Conditional Proof</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilemma Reasoning</td>
</tr>
<tr>
<td>Conditional Proof</td>
</tr>
<tr>
<td>Indirect Proof</td>
</tr>
</tbody>
</table>

Updating our lists of what we know and what we want, we get:

<table>
<thead>
<tr>
<th>What we know</th>
<th>What we want</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → R</td>
<td>[(P ∨ Q) → R]</td>
</tr>
<tr>
<td>¬R → ¬Q</td>
<td>[Q → R]</td>
</tr>
<tr>
<td>P ∨ Q</td>
<td>P</td>
</tr>
<tr>
<td>Q</td>
<td>[R]</td>
</tr>
<tr>
<td>¬R</td>
<td>¬R</td>
</tr>
<tr>
<td>Any contradiction</td>
<td></td>
</tr>
</tbody>
</table>

We've now got a point of overlap between the two lists in ¬R, so we check to see why we wanted ¬R. It turns out that we wanted it to let us move forward from ¬R → ¬Q, so we do so now:
(1) \( P \to R \)  
(2) \( \neg R \to \neg Q \)  
(3) Show \((P \lor Q) \to R\)  
(4) \( P \lor Q \)  
(5) Show \( Q \to R \)  
(6) \( Q \)  
(7) Show \( R \)  
(8) \( \neg R \)  
(9) \( \neg Q \)  
(\( \neg R \to \neg Q \) is a contradiction)  
What we want: \( R \)

<table>
<thead>
<tr>
<th>What we know</th>
<th>What we want</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P \to R )</td>
<td>((P \lor Q) \to R)</td>
</tr>
<tr>
<td>( \neg R \to \neg Q )</td>
<td>( Q \to R )</td>
</tr>
<tr>
<td>( P \lor Q )</td>
<td>( P )</td>
</tr>
<tr>
<td>( Q )</td>
<td>( R )</td>
</tr>
<tr>
<td>( \neg R )</td>
<td>( \neg R )</td>
</tr>
<tr>
<td>( \neg Q )</td>
<td>Any contradiction</td>
</tr>
</tbody>
</table>

This adds \( \neg Q \) to our wish list, yielding:

<table>
<thead>
<tr>
<th>What we know</th>
<th>What we want</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P \to R )</td>
<td>((P \lor Q) \to R)</td>
</tr>
<tr>
<td>( \neg R \to \neg Q )</td>
<td>( Q \to R )</td>
</tr>
<tr>
<td>( P \lor Q )</td>
<td>( P )</td>
</tr>
<tr>
<td>( Q )</td>
<td>( R )</td>
</tr>
<tr>
<td>( \neg R )</td>
<td>( \neg R )</td>
</tr>
<tr>
<td>( \neg Q )</td>
<td>Any contradiction</td>
</tr>
</tbody>
</table>

But now we're set – we've got a contradiction in what we know, which matches with the contradiction in what we want. So we just reiterate \( Q \) into the indirect proof. This completes both the indirect and the conditional proof:

<table>
<thead>
<tr>
<th>What we know</th>
<th>What we want</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P \to R )</td>
<td>((P \lor Q) \to R)</td>
</tr>
<tr>
<td>( \neg R \to \neg Q )</td>
<td>( Q \to R )</td>
</tr>
<tr>
<td>( P \lor Q )</td>
<td>( P )</td>
</tr>
<tr>
<td>( Q )</td>
<td>( R )</td>
</tr>
<tr>
<td>( \neg R )</td>
<td>( \neg R )</td>
</tr>
<tr>
<td>( \neg Q )</td>
<td>Any contradiction</td>
</tr>
</tbody>
</table>

Now that we've gotten \( Q \to R \), we go back and see why we wanted it. We then discover that we wanted it as half of what we need to use \( \lor E \) on \( P \lor Q \). We already had \( P \to R \), the other half of what we needed, so now we can move forward from \( P \lor Q \). Doing so will give us \( R \), which will complete the conditional proof and thereby the proof as a whole:
As you get better at doing proofs, more and more of these moves will be immediately obvious to you, and you'll need to do less and less explicit work in keeping track of what you know and what you need and how to link the two.

**Strategies for Difficult Times**

The techniques of back-and-forth proof construction and modular implementation of elements of a structural map will guide you pretty painlessly through about 90% of all proofs. However, there are some more elusive proofs where these techniques give out. In this section, I will give a few more strategic tips for tackling these more difficult proofs. I don't promise that these tips will get you through every difficulty – there may come a time when you simply have to rely on your own creativity and ingenuity – but they will guide you past a lot of hurdles.

(In a bit we will give a completely mechanical proof construction algorithm which is guaranteed to handle absolutely any proof. However, this algorithm rarely gives anywhere near the most efficient proof – it will often take a few hundred lines to complete a proof that could have been done in a dozen lines – so we won't rely on it for the actual practice of proof construction. Its interest will be strictly theoretical, as we will see.)

We've already introduced two tips for navigating through difficult spots in proofs:

- **When all else fails, try an indirect proof to get something you want.**
- **When in doubt, try to derive whatever result will be of the most help.**

These two fall-back strategies are generic ones that can be invoked any time that the standard methods aren't working. Now we will look at some techniques for two types of proofs that tend to be particularly difficult: those that require proving disjunctions and those that require moving forward from negations.

In the chart of back-and-forth strategies which I gave earlier, I recommended proving a disjunction by deriving one of its disjuncts and then using \( \lor \)I. However, the fact is that this strategy simply won't work in most cases. The problem is that if we were able to derive one of the disjuncts, then in most cases we would simply want to stop there, rather than adding on a superfluous disjunction, since that extra disjunction does nothing but weaken the conclusion. Thus, for example, we can prove:

- **P, P → R, ∴ R \lor T**

using the technique of \( \lor \)I to get the final disjunction:

<table>
<thead>
<tr>
<th>(1) P</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) P → R</td>
<td>A</td>
</tr>
<tr>
<td>(3) Show R \lor T</td>
<td></td>
</tr>
<tr>
<td>(4) R</td>
<td>\lor E, 2,1</td>
</tr>
<tr>
<td>(5) R \lor T</td>
<td>\lor E, 4</td>
</tr>
</tbody>
</table>

But we could also simply derive \( R \) (by stopping at line 4), and since, as a general rule, we're interested in deriving the strongest conclusion we can from our assumptions, we are unlikely in most natural contexts to want to derive \( R \lor T \) in this way.

Where the strategy of deriving a disjunction using \( \lor \)I from one of the disjuncts does work, it is usually because we are deriving the disjunction from an assumption which is itself a disjunction. Consider, for example, the proof of:

- **(P → Q) \lor (¬S → R), P & ¬R, ∴ Q \lor S**

which proceeds as follows:

| (1) (P → Q) \lor (¬S → R) | A   |
| (2) P & ¬R | A   |
Here we do derive \( Q \lor S \) by using \( \lor I \) on one of the disjuncts -- in fact, we do so twice. Since one of our assumptions is a disjunction, we are able first to show that the assumption disjunct \( P \rightarrow Q \) implies the conclusion disjunct \( Q \), and then use \( \lor I \) to get \( Q \lor S \), and second to show that the assumption disjunct \( \neg S \rightarrow R \) implies the conclusion disjunct \( S \), and then use \( \lor I \) to get \( Q \lor S \). Since each assumption disjunct entails a different conclusion disjunct, we can't get a stronger conclusion by dropping one of the disjuncts -- the disjunctive form of the conclusion is essential here.

This is a common pattern for proving disjunctions -- we start from a disjunction, and thus invoke dilemma reasoning to create a "two-track" proof, each track of which ends at a different disjunct of the disjunctive conclusion (and then uses \( \lor I \) to weaken that disjunct to the disjunction). However, this pattern requires that we have a disjunctive assumption to trigger the two-track process in the first place. The real difficulties with proving disjunctions arise when we have a disjunctive conclusion but no disjunctive assumption.

There are two techniques which are useful for making it through such proofs. Both require some extra work, but they are frequently the only good way to get through the proof. The first technique is to force a disjunctive assumption into the proof by deriving some sentence of the form \( \Phi \lor \neg \Phi \). We've seen earlier how such a sentence can always be derived:

\[
\begin{align*}
(1) & \text{ Show } \Phi \lor \neg \Phi \\
(2) & \text{ Show } \neg (\Phi \lor \neg \Phi) \\
(3) & \text{ Show } \neg \Phi \\
(4) & \Phi \\
(5) & \Phi \lor \neg \Phi \\
(6) & \neg (\Phi \lor \neg \Phi) \\
(7) & \Phi \lor \neg \Phi
\end{align*}
\]

(Note, by the way, that this is not a genuine proof, because \( \Phi \) is not a sentence in the language. It is rather a proof schema -- a recipe for constructing a proof given a particular choice of sentence for \( \Phi \).) Once we derive such a tautologous disjunction, we can use it to trigger a dilemma reasoning, and then (as above) reason from each disjunct to a different disjunct of the conclusion.

Here's an example of this strategy. We will prove the claim:

\[
\begin{align*}
\text{• } P &\leftrightarrow Q \therefore (P \land Q) \lor (\neg P \land \neg Q)
\end{align*}
\]

Here we are trying to prove a disjunctive conclusion with no disjunctive assumption. We can't derive either disjunct directly from \( P \leftrightarrow Q \), since \( P \leftrightarrow Q \) doesn't force \( P \) and \( Q \) both to be true, and also doesn't force \( P \) and \( Q \) both to be false (it just forces that one or the other of these will happen). So we'll start by deriving \( P \lor \neg P \), and then getting \( P \land Q \) from \( P \) and \( \neg P \land \neg Q \) from \( \neg P \). The proof thus proceeds as follows:

\[
\begin{align*}
(1) & P \leftrightarrow Q \\
(2) & \text{ Show } (P \land Q) \lor (\neg P \land \neg Q) \\
(3) & \text{ Show } P \lor \neg P \\
(4) & \neg (P \lor \neg P) \\
(5) & \text{ Show } \neg P \\
(6) & P \\
(7) & \text{ Show } \neg P \\
(8) & \neg (P \lor \neg P)
\end{align*}
\]
Once we've derived $P \lor \neg P$ and used it to trigger a dilemma reasoning, the rest of the proof flows quite easily, and what promised to be a very challenging proof falls into place immediately.

The other useful technique for proving a disjunction in the absence of a disjunctive assumption is to derive instead a conditional and then convert it into a disjunction. Suppose that I'm trying to derive $\Phi \lor \Theta$. I will instead derive $\neg \Phi \rightarrow \Theta$, and then convert $\neg \Phi \rightarrow \Theta$ into $\Phi \lor \Theta$ using the following technique:

1. $\neg \Phi \rightarrow \Theta \quad A$
2. Show $\Phi \lor \Theta$
3. $\neg (\Phi \lor \Theta) \quad AIP$
4. Show $\Phi$
5. $\neg \Phi \quad AIP$
6. $\Theta \quad \rightarrow E, 1,5$
7. $\Phi \lor \Theta \quad \lor E, 6$
8. $\neg (\Phi \lor \Theta) \quad R, 3$
9. $\Phi \lor \Theta \quad \lor E, 4$

This technique shifts our goal from a disjunction to a conditional, and thus allows us to use the familiar technique of proving a conditional via conditional proof.

Let's look at an example of this technique. We will prove the following result:
- $\neg Q \rightarrow R, \therefore \neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q)$
We will do so by first proving:
- $\neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q)$
and then converting it into:
- $\neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q)$
using the technique we just sketched. The proof proceeds as follows:

1. $\neg Q \rightarrow R \quad A$
2. Show $\neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q)$
3. $\neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q) \quad AIP$
4. Show $\neg (P \leftrightarrow \neg R) \lor (P \rightarrow Q)$
5. $\neg (P \leftrightarrow \neg R) \quad ACP$
6. Show $P \rightarrow Q$
7. $P \quad ACP$
8. Show $Q$
9. $\neg Q \quad AIP$
10. $R \quad \rightarrow E, 1,9$
11. $P \leftrightarrow \neg R \quad \leftrightarrow E, 5$
12. $\neg R \quad \leftrightarrow E, 11,7$
13. Show $\neg (P \leftrightarrow \neg R)$
14. $\neg (P \leftrightarrow \neg R) \quad AIP$
Lines 4 through 12 of this proof derive the conditional, and then lines 13 through 19 convert the conditional into the desired disjunction.

We now have two emergency techniques for dealing with proofs ending in disjunctions, when those proofs give us trouble. Next I want to look at the other major category of difficult proofs -- those starting from negations. The difficulty with a proof which has a negation as an assumption (or which at some point has a negation as one of the things that you know) is that our closest approximation to a rule of negation elimination is $\neg\neg$, which is only useful if the negation in question is actually a double negation.

What we are forced to do, then, to complete a proof from a negated disjunction, is to keep waiting in hopes that the negated sentence will at some point show up on the list of things that we want, thereby allowing us to meet in the middle and move forward. Sometimes this happens quite simply, as in the proof of:

- $(P \land Q), P \to (\neg (P \land Q) \to R), \therefore P \to R$

which proceeds as follows:

<table>
<thead>
<tr>
<th>(1) $\neg (P \land Q)$</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) $P \to (\neg (P \land Q) \to R)$</td>
<td>A</td>
</tr>
<tr>
<td>• (3) Show $P \to R$</td>
<td></td>
</tr>
<tr>
<td>(4) $P$</td>
<td>ACP</td>
</tr>
<tr>
<td>(5) $(P \land Q) \to R$</td>
<td>$\to E$, 2, 4</td>
</tr>
<tr>
<td>(6) R</td>
<td>$\to E$, 5, 1</td>
</tr>
</tbody>
</table>

The appearance of the conditional $(P \land Q) \to R$ on line 5 prompts us to add its antecedent -- $(P \land Q)$ -- to the list of things we want, so that we can use $\to E$. $(P \land Q)$ is already on the list of things we know (since it is one of our starting assumptions), and this convergence allows us to move forward with the proof.

Unfortunately, things often don’t work out so conveniently. One of the more frequent ways for negated assumptions to come into use in a proof is as part of a contradiction for completing an indirect proof, and this pattern can be harder to spot, since we don’t know exactly what contradiction we are looking for. Here’s an example. We will prove:

- $(P \to \neg P), \therefore P$

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>(1) $P \to \neg P$</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Show $P$</td>
<td>ACP</td>
</tr>
<tr>
<td>(3) $\neg P$</td>
<td>AIP</td>
</tr>
<tr>
<td>(4) Show $P \to \neg P$</td>
<td></td>
</tr>
<tr>
<td>(5) $P$</td>
<td>ACP</td>
</tr>
<tr>
<td>(6) $\neg P$</td>
<td>R, 3</td>
</tr>
<tr>
<td>(7) $(P \to \neg P)$</td>
<td>R, 1</td>
</tr>
</tbody>
</table>

In this proof, the contradiction which completes the indirect proof of $P$ is between $P \to \neg P$ and $(P \to \neg P)$. Had we not known to be on the lookout for a contradiction involving $(P \to \neg P)$, it is unlikely that we would have considered deriving $P \to \neg P$, and the proof would have stalled.

The previous example was a relatively simple case. In more complex cases, the use of the negated assumption can be buried a few levels deep in indirect proofs, and it can be quite challenging correctly to construct the proof. Consider, for example, the proof of:

- $(P \land Q), \therefore P \lor \neg Q$

The proof proceeds as follows:

<table>
<thead>
<tr>
<th>(1) $(P \land Q)$</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Show $P \lor \neg Q$</td>
<td></td>
</tr>
<tr>
<td>(3) $\neg P \lor \neg Q$</td>
<td>AIP</td>
</tr>
<tr>
<td>(4) Show $\neg P$</td>
<td></td>
</tr>
<tr>
<td>(5) $P$</td>
<td>AIP</td>
</tr>
</tbody>
</table>
In this proof, we have three nested levels of indirect proof, and it is not until we get to the third level that we are ready to use the initial negated assumption as half of a contradiction.

The general advice, then, for getting through difficult proofs which involve moving forward from a negation, is this:

• Keep starting indirect proofs, until you accumulate enough information via the various AIPs to form a contradiction with the negation you already know.

Another useful technique for moving forward from negations is to try to force the negation in a step. Suppose that among what I know is:

• \( \neg (P \leftrightarrow Q) \)

We know from our earlier study of equivalences in sentential logic that \( \neg (P \leftrightarrow Q) \) is equivalent to \( P \leftrightarrow \neg Q \), so there should be some proof which leads from \( \neg (P \leftrightarrow Q) \) to \( P \leftrightarrow \neg Q \). If we can construct that proof, then we can replace the negation on our list of what we know with the biconditional, and we’ve got techniques in place for moving forward from biconditionals. There are similar distribution equivalences which push inward negations past each type of connective:

• \( \neg (P \& Q) \) is equivalent to \( \neg P \lor \neg Q \)
• \( \neg (P \lor Q) \) is equivalent to \( \neg P \& \neg Q \)
• \( \neg (P \rightarrow Q) \) is equivalent to \( P \& \neg Q \)
• \( \neg (P \leftrightarrow Q) \) is equivalent to \( P \leftrightarrow \neg Q \)

So given any negation among what we know, if we can manage to derive the appropriate inference, we can get rid of that negation (really, delay dealing with it temporarily, since a new negation always emerges lower down in the sentence after the distribution). All of these equivalences are in fact derivable, so this is a viable tactic. However, I’m not going to go through the techniques for deriving each, because we’ll be tackling that issue in more detail in the next section.

Derived Proof Rules

Suppose I'm asked to prove the following result:

\[ P \rightarrow Q, \neg P \rightarrow \neg S, R \rightarrow S, \neg R \rightarrow \neg T, U \rightarrow T, \neg Q \therefore \neg U \]

The proof is quite straightforward to form, if somewhat lengthy. It proceeds as follows:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>( P \rightarrow Q )</td>
<td>A</td>
</tr>
<tr>
<td>(2)</td>
<td>( \neg P \rightarrow \neg S )</td>
<td>A</td>
</tr>
<tr>
<td>(3)</td>
<td>( R \rightarrow S )</td>
<td>A</td>
</tr>
<tr>
<td>(4)</td>
<td>( \neg R \rightarrow \neg T )</td>
<td>A</td>
</tr>
<tr>
<td>(5)</td>
<td>( U \rightarrow T )</td>
<td>A</td>
</tr>
<tr>
<td>(6)</td>
<td>( \neg Q )</td>
<td>A</td>
</tr>
<tr>
<td>(7)</td>
<td>Show ( \neg U )</td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>Show ( \neg P )</td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>( P )</td>
<td>AIP</td>
</tr>
<tr>
<td>(10)</td>
<td>( \neg Q )</td>
<td>( \neg E, 1,9 )</td>
</tr>
<tr>
<td>(11)</td>
<td>( \neg Q )</td>
<td>R, 6</td>
</tr>
<tr>
<td>(12)</td>
<td>( \neg S )</td>
<td>( \neg E, 2,8 )</td>
</tr>
<tr>
<td>(13)</td>
<td>Show ( \neg R )</td>
<td></td>
</tr>
<tr>
<td>(14)</td>
<td>( R )</td>
<td>AIP</td>
</tr>
<tr>
<td>(15)</td>
<td>( S )</td>
<td>( \neg E, 3,14 )</td>
</tr>
<tr>
<td>(16)</td>
<td>( \neg S )</td>
<td>R, 12</td>
</tr>
<tr>
<td>(17)</td>
<td>( \neg T )</td>
<td>( \neg E, 4,13 )</td>
</tr>
<tr>
<td>(18)</td>
<td>Show ( \neg U )</td>
<td></td>
</tr>
<tr>
<td>(19)</td>
<td>( U )</td>
<td>AIP</td>
</tr>
<tr>
<td>(20)</td>
<td>( T )</td>
<td>( \neg E, 5,18 )</td>
</tr>
<tr>
<td>(21)</td>
<td>( \neg T )</td>
<td>R, 17</td>
</tr>
</tbody>
</table>
In this proof, we three times run through a short routine which derives \( \neg \Phi \) from \( \Phi \rightarrow \Theta \) and \( \neg \Theta \) -- that is, which derives the falsity of the antecedent from the truth of the conditional and the falsity of the consequent. In each case, the routine has the same structure. We:

- Begin an indirect proof of \( \neg \Phi \)
- Take \( \Phi \) as AIP
- Combine \( \Phi \) with \( \Phi \rightarrow \Theta \) to produce \( \Theta \)
- Reiterate \( \neg \Theta \) to form a contradiction.

The choice of \( \Phi \) and \( \Theta \) varies, but what we do with them is in each case the same.

It thus becomes natural to wonder: if I know how to derive \( \neg \Phi \) from \( \Phi \rightarrow \Theta \) and \( \neg \Theta \), why do I have to keep doing the work over and over again? And, of course, this is hardly the only case like this. Already in the proofs we've done we've seen several routines which occur again and again. We've needed a few times to introduce \( \Phi \lor \neg \Phi \), for example, and deriving the truth of a conditional from the falsity of its antecedent has been a recurring theme. To avoid the need for such repetition in our proofs, we are now going to introduce the idea of a derived proof rule.

A derived proof rule can be thought of as an abbreviation for several steps in a proof. It can also be thought of as akin to a program routine, which through a single call performs a whole sequence of operations. The idea is that once we show that we can make a particular kind of transition using our basic proof rules -- such as the transition from \( \Phi \rightarrow \Theta \) and \( \neg \Theta \) to \( \neg \Phi \) -- we can encode that transition in a new proof rule, thereby allowing us from then on to make it in a single step, instead of going through the same procedure over and over again. So, for example, we will now introduce the derived rule of \( \rightarrow \text{E}^* \) (which we will sometimes call alternative conditional elimination, and which historically has had the name modus tollens, or -- for the particularly pretentious -- modus tollendo tollens), which has the following input-output structure:

\[
\Phi \rightarrow \Theta \\
\neg \Theta \\
\hline
\neg \Phi 
\]

It's important to realize that derived proof rules cannot be introduced arbitrarily. When we introduce a derived proof rule, it is not supposed to add any new power to the proof system, but is simply supposed to give us a shorter way of doing something that we could have done already with the core rules. Therefore each derived rule requires a justification. The required justification is a recipe showing that the move from the inputs of the derived rule to the outputs of the derived rule can always be accomplished using the core proof rules. In the case of \( \rightarrow \text{E}^* \), the justification is:

\[
\begin{array}{l}
(1) \quad \Phi \rightarrow \Theta \\
(2) \quad \neg \Theta \\
(3) \text{Show } \neg \Phi \\
(4) \quad \Phi \\
(5) \quad \Theta \\
(6) \quad \neg \Theta \\
(7) \text{Show } \neg \Theta \\
(8) \quad \neg P \\
(9) \quad \neg S \\
(10) \quad \neg R \\
\end{array}
\]

Since we can take \( \Phi \) and \( \Theta \) to be any sentences we want without disturbing the validity of this reasoning, we know that any occurrence of \( \rightarrow \text{E}^* \) can be expanded out into an indirect proof in this manner, using only the core rules.

Adding \( \rightarrow \text{E}^* \) allow us to save quite a bit of time in our proofs, since we no longer need to take a circuitous route to extracting the negation of the antecedent from a conditional and the negation of the consequent. Consider how the proof of:

- \( P \rightarrow Q \), \( \neg P \rightarrow \neg S \), \( R \rightarrow S \), \( \neg R \rightarrow \neg T \), \( U \rightarrow T \), \( \neg Q \) \: \: \neg U

goes with \( \rightarrow \text{E}^* \):

\[
\begin{array}{l}
(1) \quad P \rightarrow Q \\
(2) \quad \neg P \rightarrow \neg S \\
(3) \quad R \rightarrow S \\
(4) \quad \neg R \rightarrow \neg T \\
(5) \quad U \rightarrow T \\
(6) \quad \neg Q \\
(7) \text{Show } \neg U \\
(8) \quad \neg P \\
(9) \quad \neg S \\
(10) \quad \neg R \\
\end{array}
\]

Since we can take \( \Phi \) and \( \Theta \) to be any sentences we want without disturbing the validity of this reasoning, we know that any occurrence of \( \rightarrow \text{E}^* \) can be expanded out into an indirect proof in this manner, using only the core rules.
The proof with \( \rightarrow E^* \) is nine lines shorter than the original proof, and the substantive portion of the proof (that is, that portion after the initial assumptions and Show line) is about a third the length of the substantive portion of the original proof.

The general plan, then, is that any time we find a pattern of inference which is repeatedly useful, we will see if we can isolate that inference plan and construct a derived rule out of it. In this way we will eventually build up a library of derived rules which act as time-saving devices, making proofs easier to construct and inspect. This is a completely open-ended process -- there is no limit to the number of derived rules we can introduce. In this section I will set out what in my experience are the most useful derived rules, but there's nothing sacred about the list I will give. It can be expanded arbitrarily. Importantly, none of these derived rules adds anything genuinely new to the proof system. If we add 500 derived rules, we will at the end of the day be able to prove exactly the same results that we could prove before introducing the derived rules (just more quickly), because all we are doing is giving ourselves ways to abbreviate portions of proofs.

We're now going to go through a list of derived rules. For each one, I will introduce the rule and give a justification of it. The derived rules will fall into four major categories:

- Rearrangement rules
- Distribution rules
- Rules for interconverting connectives
- Miscellaneous additional rules

The rearrangement rules are rules which allow you to take a single sentence and rearrange some of its internal structure. Perhaps the easiest of these rules is that of conjunction commutivity (\( \&C \)), which has the following input-output pattern:

\[
\&C: \quad \Phi \& \Theta \quad \Rightarrow \quad \Theta \& \Phi
\]

Recall that the double line between the input and the output indicates that this is a bidirectional rule -- you can infer from \( \Phi \& \Theta \) to \( \Theta \& \Phi \), or from \( \Theta \& \Phi \) to \( \Phi \& \Theta \). Thus the derived rule will require two justifications, one for each direction. The top-to-bottom direction is justified as follows:

\[
\text{(1) } \Phi \& \Theta \quad \text{A} \\
\text{(2) Show } \Theta \& \Phi \\
\text{(3) } \Theta \\
\text{(4) } \Phi \\
\text{(5) } \Theta \& \Phi \quad \&I, 4, 3
\]

The top-to-bottom justification is then essentially identical:

\[
\text{(1) } \Theta \& \Phi \\
\text{(2) Show } \Phi \& \Theta \\
\text{(3) } \Phi \\
\text{(4) } \Theta \\
\text{(5) } \Phi \& \Theta \quad \&I, 4, 3
\]

From now on, when the two justificatory directions of a bidirectional derived rule are essentially identical (when one can be gotten from the other just by swapping metavariables), we will only give one direction. The rule of \( \&C \) captures the fact that the ordering of the conjuncts in a conjunction does not matter, and allows us to rearrange that ordering in one step, without going through the work of disassembling and reconstructing the conjunction.

**Question:** Why does \( \&C \) need to be a bidirectional rule? Can't you infer in both directions just by changing your choice of \( \Phi \) and \( \Theta \)? To go from \( \Phi \& \Theta \) to \( \Theta \& \Phi \), let \( \Phi \) be \( P \) and \( \Theta \) be \( Q \); to go from \( \Theta \& \Phi \) to \( \Phi \& \Theta \), let \( \Phi \) be \( Q \) and \( \Theta \) be \( P \).

**Answer:** This is quite correct -- making \( \&C \) a bidirectional rule adds nothing to its inferential power. However, when we come to the concept of Replacement below, it will be important which rules are bidirectional and which are monodirectional, and it will turn out to be useful to make as many rules as possible bidirectional.
Our next derived rule is that of disjunction commutivity, or $\lor^C$. Disjunctions, like conjunctions, are commutative – you can put the disjuncts in any order without changing the truth conditions of the sentence. So we want a rule which allows us to swap the order of disjuncts, and which thus has the following input-output pattern:

$\lor^C: \begin{align*}
\Phi \lor \Theta \\
\Theta \lor \Phi
\end{align*}$

$\lor^C$, like $\&^C$, is a bidirectional rule. We have already given a proof of:

- $P \lor Q, \therefore Q \lor P$

and since this proof, with $P$ replaced by $\Phi$ and $Q$ replaced by $\Theta$, serves as a justification for $\lor^C$, I won’t give a justification here.

Biconditionals are also commutative, since $P \leftrightarrow Q$ is equivalent to $Q \leftrightarrow P$, so we can also have a rule of biconditional commutivity, or $\leftrightarrow^C$. This rule has the following input-output pattern:

$\leftrightarrow^C: \begin{align*}
\Phi \leftrightarrow \Theta \\
\Theta \leftrightarrow \Phi
\end{align*}$

To justify the rule of $\leftrightarrow^C$, we need to derive $\Theta \leftrightarrow \Phi$ from $\Phi \leftrightarrow \Theta$ (as I mentioned above, we technically need a justification in the other direction as well, since it’s a bidirectional rule. But this justification will be structurally identical to the one I’m about to give, so I’ll omit it.). The proof proceeds as follows:

| (1) $\Theta \leftrightarrow \Theta$ | A |
| (2) Show $\Theta \leftrightarrow \Theta$ | |
| (3) Show $\Theta \rightarrow \Phi$ | |
| (4) $\Theta$ | ACP |
| (5) $\Phi$ | $\leftrightarrow^E, 1,4$ |
| (6) Show $\Phi \rightarrow \Theta$ | |
| (7) $\Phi$ | ACP |
| (8) $\Theta$ | $\leftrightarrow^E, 17,7$ |

Each application of $\leftrightarrow^C$ thus saves six steps in a proof.

Should we also have a rule of conditional commutivity to go with the other three? Only if conditionals are in fact commutative – that is, if $P \rightarrow Q$ is equivalent to $Q \rightarrow P$. Of course, it is not, so we should not have a rule. If we look at the truth tables for the four binary connectives side-by-side, we can easily see how commutivity fails for the conditional:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P &amp; Q$</th>
<th>$P \lor Q$</th>
<th>$P \rightarrow Q$</th>
<th>$P \leftrightarrow Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Swapping $P$ and $Q$ makes a difference only on the middle two rows (on the top row, we are just exchanging a $T$ for a $T$, and on the bottom an $F$ for an $F$), so a connective will be commutative just in case it produces the same output for both of the middle two rows. $\&$, $\lor$, and $\leftrightarrow$ all satisfy this condition, but $\rightarrow$ does not, and hence is not commutative.

In addition to commutivity properties, some of the connectives have interesting associativity properties. While commutivity tells us that the order of the sentences connected by the connective doesn't matter, associativity tells us that the grouping of the sentences joined by the connective doesn't matter. So, for example, $\&$ is associative, because the following two sentences are equivalent:

- $(P \& Q) \& R$
- $P \& (Q \& R)$

It doesn't matter whether we first connect $P$ with $Q$ using $\&$, and then connect the resulting $P \& Q$ to $R$ with $\&$, or first connect $Q$ with $R$, and then connect the resulting $Q \& R$ with $P$. We thus want a rule of conjunction associativity, or $\&^A$, which will have the following input-output pattern:

$(\Phi \& \Theta) \& \Psi$
Like all of the commutivity rules, this rule is bidirectional. However, the two directions of justification are not structurally
identical (although they are quite similar), so we will give both. The first is:

\[(1) (\Phi \& \Theta) \& \Psi \quad \text{A} \]
\[(2) \text{Show } \Phi \& (\Theta \& \Psi) \]
\[(3) \Psi \quad \&E, 1 \]
\[(4) \Phi \& \Theta \quad \&E, 1 \]
\[(5) \Phi \quad \&E, 4 \]
\[(6) \Theta \quad \&E, 4 \]
\[(7) \Theta \& \Psi \quad \&I, 6,3 \]
\[(8) \Phi \& (\Theta \& \Psi) \quad \&I, 5,7 \]

The justification of the bottom-to-top direction is:

\[(1) \Phi \& (\Theta \& \Psi) \quad \text{A} \]
\[(2) \text{Show } (\Phi \& \Theta) \& \Psi \]
\[(3) \Theta \& \Psi \quad \&E, 2 \]
\[(4) \Theta \quad \&E, 3 \]
\[(5) \Psi \quad \&E, 3 \]
\[(6) \Phi \quad \&E, 2 \]
\[(7) \Phi \& \Theta \quad \&I, 6,4 \]
\[(8) (\Phi \& \Theta) \& \Psi \quad \&I, 6,5 \]

\textbf{A Fussy Point:} I said earlier that derived rules are to be thought of as abbreviations for proof fragments carried
out under the core rules. But note that there is frequently more than one way to justify a given derived rule. In the
above derivation of the bottom-to-top direction of \&A, for example, I could have extracted \(\Phi\) from \(\Phi \& (\Theta \& \Psi)\) first,
rather than \(\Theta \& \Psi\). So which of these multiple justifications does the rule abbreviate? Just so that the
abbreviations have determinate content, I will stipulate that each derived rule abbreviates exactly the justificatory
recipe that I give here. Having it abbreviate some other recipe that also achieved the same result wouldn't cause
anything to go wrong – this is purely a matter of bookkeeping, to make sure everything is precisely defined.

Disjunctions are also associative, so we can have a rule of disjunction associativity, or \(\vee C\). This rule will have the following
input-output pattern:

\[\text{\&A: } \Phi \vee (\Theta \vee \Psi) \]
\[\vdash \quad \text{\&A: } \Phi \vee (\Theta \vee \Psi) \]

We have \textbf{already seen} a justification of this rule in the bottom-to-top direction, so here we will only justify the top-to-bottom
direction:

\[(1) \Phi \vee (\Theta \vee \Psi) \quad \text{A} \]
\[(2) \text{Show } (\Phi \vee \Theta) \vee \Psi \]
\[(3) \text{Show } \Phi \rightarrow ((\Phi \vee \Theta) \vee \Psi) \]
\[(4) \Phi \quad \text{ACP} \]
\[(5) \Phi \vee \Theta \quad \text{vl, 4} \]
\[(6) (\Phi \vee \Theta) \vee \Psi \quad \text{vl, 5} \]
\[(7) \text{Show } (\Theta \vee \Psi) \rightarrow ((\Phi \vee \Theta) \vee \Psi) \]
\[(8) \Theta \vee \Psi \quad \text{ACP} \]
\[(9) \text{Show } \Theta \rightarrow ((\Phi \vee \Theta) \vee \Psi) \]
\[(10) \Theta \text{ACP} \]
\[(11) \Phi \vee \Theta \quad \text{vl, 10} \]
\[(12) (\Phi \vee \Theta) \vee \Psi \quad \text{vl, 11} \]
\[(13) \text{Show } \Psi \rightarrow ((\Phi \vee \Theta) \vee \Psi) \]
\[(14) \Psi \quad \text{ACP} \]
\[(15) (\Phi \vee \Theta) \vee \Psi \quad \text{vl, 14} \]
Biconditionals are also associative, so we need a rule of $\leftrightarrow A$ with the following input-output pattern:

$$\leftrightarrow A: \quad \Phi \leftrightarrow (\Theta \leftrightarrow \Psi)$$

This rule has a formidable justification in each direction. For this reason, I am going to defer giving it until we have in place a few other derived rules that can shorten the process.

Conditionals, however, are not associative. We can test this easily using a truth table. Comparing $(P \rightarrow Q) \rightarrow R$ with $P \rightarrow (Q \rightarrow R)$, we get:

<table>
<thead>
<tr>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>$P \rightarrow Q$</th>
<th>$P \rightarrow (Q \rightarrow R)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

The two rows highlighted in red give different results for the two sentences, thus showing that they do not imply each other. (Actually, we see here that $(P \rightarrow Q) \rightarrow R$ implies $P \rightarrow (Q \rightarrow R)$, but not the other way around. We could therefore set up a monodirectional derived rule corresponding to that implication, but the pattern won't come up often enough for it to be worth it).

Neither commutivity nor associativity have applied to conditionals, but we will set out two different rearrangement rules for conditionals. The first of these – which we will call $\rightarrow \rightarrow$ – allows us to take two antecedents in a chained conditional and combine them into a single conjunctive antecedent. It has the following input-output pattern:

$$\Phi \rightarrow (\Theta \rightarrow \Psi)$$

In the top-to-bottom direction, the rule is justified as follows:

1. $\Phi \rightarrow (\Theta \rightarrow \Psi)$  
2. Show $\Phi \& \Theta$  
3. $\Phi$  
4. $\Phi$  
5. $\Theta$  
6. $\Theta$  
7. $\Psi$  

In the bottom-to-top direction, the rule is justified as follows:

1. $(\Phi \& \Theta) \rightarrow \Psi$  
2. Show $\Phi \rightarrow (\Theta \rightarrow \Psi)$  
3. $\Phi$  
4. $\Phi$  
5. $\Theta$  
6. $\Theta$  
7. $\Psi$  

The second rearrangement rule for the conditional allows us to swap the order of two chained antecedents. The resulting rule of conditional swap, or $\rightarrow \rightarrow Sw$, thus has the following input-output pattern:
\[ \Phi \to (\Theta \to \Psi) \]

We only need one direction of justification (since the two directions are structurally identical), so we need only give:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>( \Phi \to (\Theta \to \Psi) )</td>
</tr>
<tr>
<td>(2)</td>
<td>Show ( \Theta \to (\Phi \to \Psi) )</td>
</tr>
<tr>
<td>(3)</td>
<td>( \Theta )</td>
</tr>
<tr>
<td>(4)</td>
<td>Show ( \Phi \to \Psi )</td>
</tr>
<tr>
<td>(5)</td>
<td>( \Phi )</td>
</tr>
<tr>
<td>(6)</td>
<td>( \Theta \to \Psi )</td>
</tr>
<tr>
<td>(7)</td>
<td>( \Psi )</td>
</tr>
</tbody>
</table>

This brings us to the end of the rearrangement rules. Next, we will look at a collection of distribution rules, which are designed to interchange the scopes of two different connectives in a sentence. First, we will give a sequence of four rules, each of which allows a negation to be moved from a wide-scope position in front of some other connective to a narrow-scope position.

First we have the rule of negation-conjunction interchange, or \( \neg \& \), which has the following input-output pattern:

\[ \neg (\Phi \& \Theta) \]

We've already seen the top-to-bottom direction of justification, so here we'll just give the bottom-to-top direction:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>( \neg \Phi \vee \neg \Theta )</td>
</tr>
<tr>
<td>(2)</td>
<td>Show ( \neg (\Phi &amp; \Theta) )</td>
</tr>
<tr>
<td>(3)</td>
<td>Show ( \neg \Phi \to \neg(\Phi &amp; \Theta) )</td>
</tr>
<tr>
<td>(4)</td>
<td>( \neg \Phi )</td>
</tr>
<tr>
<td>(5)</td>
<td>Show ( \neg(\Phi &amp; \Theta) )</td>
</tr>
<tr>
<td>(6)</td>
<td>( \Phi &amp; \Theta )</td>
</tr>
<tr>
<td>(7)</td>
<td>( \Phi )</td>
</tr>
<tr>
<td>(8)</td>
<td>( \neg \Phi )</td>
</tr>
<tr>
<td>(9)</td>
<td>Show ( \neg \Theta \to \neg(\Phi &amp; \Theta) )</td>
</tr>
<tr>
<td>(10)</td>
<td>( \neg \Theta )</td>
</tr>
<tr>
<td>(11)</td>
<td>Show ( \neg(\Phi &amp; \Theta) )</td>
</tr>
<tr>
<td>(12)</td>
<td>( \Phi &amp; \Theta )</td>
</tr>
<tr>
<td>(13)</td>
<td>( \Theta )</td>
</tr>
<tr>
<td>(14)</td>
<td>( \neg \Theta )</td>
</tr>
<tr>
<td>(15)</td>
<td>( \neg(\Phi &amp; \Theta) )</td>
</tr>
</tbody>
</table>

We also have a rule of negation-disjunction interchange, or \( \neg \vee \), which has the following input-output pattern:

\[ \neg(\Phi \vee \Theta) \]

\( \neg \& \) and \( \neg \vee \) together capture the DeMorgan equivalences, which give methods of distributing negations over conjunctions and disjunctions.

\( \neg \vee \) requires justification in both directions. In the top-to-bottom direction, the justification is:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>( \neg(\Phi \vee \Theta) )</td>
</tr>
<tr>
<td>(2)</td>
<td>Show ( \neg \Phi &amp; \neg \Theta )</td>
</tr>
<tr>
<td>(3)</td>
<td>Show ( \neg \Phi )</td>
</tr>
</tbody>
</table>
Negations can also be distributed over conditionals and biconditionals. For conditionals, the resulting rule of \( \neg \rightarrow \) has the following input-output pattern:

\[
\neg \rightarrow: \quad \equiv \quad \Phi \& \neg \Theta
\]

This rule requires justification in both directions. In the top-to-bottom direction, the justification is:

| (1) \( \neg (\Phi \rightarrow \Theta) \) | A |
| (2) Show \( \Phi \& \neg \Theta \) | |
| (3) Show \( \Phi \) | |
| (4) \( \neg \Phi \) | AIP |
| (5) Show \( \Phi \rightarrow \Theta \) | |
| (6) Show \( \Theta \) | |
| | (7) Show \( \Phi \rightarrow \Theta \) |
| | (8) \( \neg \Theta \) | AIP |
| | (9) \( \Phi \) | R, 6 |
| | (10) \( \neg \Phi \) | R, 4 |
| | (11) \( \neg (\Phi \rightarrow \Theta) \) | R, 1 |

In the bottom-to-top direction, the justification is:

| (1) \( \Phi \& \neg \Theta \) | A |
| (2) Show \( \neg \Theta \) | |
| (3) Show \( \Theta \) | |
| (4) \( \Phi \) | AIP |
| (5) Show \( \Phi \rightarrow \Theta \) | |
| | (6) \( \Theta \) | |
| | (7) Show \( \Theta \) | |
| | | (8) \( \Phi \) | |
| | | (9) \( \neg \Theta \) | AIP |
| | | (10) \( \neg \Phi \) | R, 4 |
| | | (11) \( \neg (\Phi \rightarrow \Theta) \) | R, 1 |

| (12) Show \( \Phi \& \neg \Theta \) | |
| (13) Show \( \Phi \rightarrow \Theta \) | |
| (14) \( \Phi \) | ACP |
| | (15) \( \Theta \) | |
| | | (16) \( \Theta \) | R, 13 |
| | | | (17) \( \Phi \& \neg \Theta \) | \&I, 3,12 |
The distribution rule for negation over the biconditional, or $\neg \leftrightarrow$, has the following input-output pattern:

$$
\neg (\Phi \leftrightarrow \Theta) \\
\equiv \equiv \equiv \equiv \\
(\Phi \leftrightarrow \neg \Theta)
$$

We have already given the top-to-bottom justification for this rule. The bottom-to-top justification is:

(1) $\Phi \leftrightarrow \neg \Theta$  
(2) Show $\neg (\Phi \leftrightarrow \Theta)$  
(3) $\Phi \leftrightarrow \Theta$  
(4) Show $\Phi \lor \neg \Phi$  
(5) $\neg (\Phi \lor \neg \Phi)$  
(6) Show $\neg \Phi$  
(7) $\Phi$  
(8) $\Phi \lor \neg \Phi$  
(9) $\neg (\Phi \lor \neg \Phi)$  
(10) $\neg \Phi$  
(11) Show $\Phi \rightarrow (\Theta \land \neg \Theta)$  
(12) $\Phi$  
(13) $\Theta$  
(14) $\rightarrow \Theta$  
(15) $\Theta \land \neg \Theta$  
(16) Show $\neg \Phi \rightarrow (\Theta \land \neg \Theta)$  
(17) $\neg \Phi$  
(18) Show $\Theta$  
(19) $\neg \Theta$  
(20) $\Phi$  
(21) $\neg \Phi$  
(22) Show $\neg \Theta$  
(23) $\Theta$  
(24) $\Phi$  
(25) $\neg \Phi$  
(26) $\Theta \land \neg \Theta$  
(27) $\Theta \land \neg \Theta$  
(28) $\Theta$  
(29) $\neg \Theta$

Question: What's going on with the derivation of $\Phi \lor \neg \Phi$ in lines 4 through 10? What does it have to do with what we're trying to prove?

Answer: This is an application of one of our strategies for difficult times -- try deriving a "free" disjunction, and then using a dilemma reasoning. $\Phi \lor \neg \Phi$ turns out to be quite useful, since $\Phi$ allows us to use both of the biconditionals, and $\neg \Phi$ allows us (somewhat more complexly) to use both of the biconditionals backward.

Question: OK, once you decided to do a dilemma reasoning from $\Phi \lor \neg \Phi$, what made you decide $\Theta \land \neg \Theta$ should be the target of it? That seems to come out of nowhere.

Answer: It's important to keep in mind here that we are doing the dilemma reasoning inside an indirect proof, and hence that our primary goal is to derive a contradiction. Since we're doing a dilemma reasoning, we need to derive a contradiction -- and the same contradiction -- from each disjunct. $\Theta \land \neg \Theta$ acts as a package disjunction, which we can extract from each wing of the dilemma reasoning, and then unpack into an explicit contradiction after we use $\lor E$. 
This gives us rules for distributing negations over each of the four binary connectives. These rules can be quite useful when trying to prove claims from negated assumptions -- by using a distribution rule, some other connective can be uncovered as the main connective, and then an appropriate strategy can be brought to bear on that new connective. Next we will give three distribution rules designed for the interaction of \& and \lor.

The first rule -- that of \lor\& -- distributes a disjunction over a conjunction to create a conjunction of disjunctions. It has the following input-output pattern:

\[ \Phi \lor (\Theta \& \psi) \]
\[ \overset{\text{=}}{=} \]
\[ (\Phi \lor \Theta) \& (\Phi \lor \psi) \]

This rule is very similar to the distribution law for addition and multiplication:

\[ x(y + z) = xy + xz \]

but with multiplication replaced by disjunction and addition replaced by conjunction.

The derived rule is bidirectional, and hence requires justification in both directions. Top-to-bottom, the justification is:

| 1 | \( \Phi \lor (\Theta \& \psi) \) | A |
| 2 | Show \( (\Phi \lor \Theta) \& (\Phi \lor \psi) \) |
| 3 | Show \( \Phi \rightarrow ((\Phi \lor \Theta) \& (\Phi \lor \psi)) \) |
| 4 | \( \Phi \) | ACP |
| 5 | \( \Phi \lor \Theta \) | \lor I, 4 |
| 6 | \( \Phi \lor \psi \) | \lor I, 4 |
| 7 | \( (\Phi \lor \Theta) \& (\Phi \lor \psi) \) | \& I, 5, 6 |
| 8 | Show \( (\Theta \& \psi) \rightarrow ((\Phi \lor \Theta) \& (\Phi \lor \psi)) \) |
| 9 | \( \Theta \& \psi \) | ACP |
| 10 | \( \Theta \) | \& E, 9 |
| 11 | \( \Phi \lor \Theta \) | \lor I, 10 |
| 12 | \( \psi \) | \& E, 9 |
| 13 | \( (\Phi \lor \Theta) \& (\Phi \lor \psi) \) | \& I, 11, 13 |
| 14 | \( (\Phi \lor \Theta) \& (\Phi \lor \psi) \) | \lor E, 1, 3, 8 |

Bottom-to-top, the justification is:

| 1 | \( (\Phi \lor \Theta) \& (\Phi \lor \psi) \) | A |
| 2 | Show \( \Phi \lor (\Theta \& \psi) \) |
| 3 | Show \( \Phi \lor \neg \Phi \) |
| 4 | \( \neg (\Phi \lor \neg \Phi) \) | AIP |
| 5 | Show \( \neg \Phi \) |
| 6 | \( \Phi \lor \neg \Phi \) | AIP |
| 7 | \( \neg (\Phi \lor \neg \Phi) \) | \lor I, 6 |
| 8 | \( \neg (\Phi \lor \neg \Phi) \) | \lor I, 5 |
| 9 | \( \Phi \lor (\Theta \& \psi) \) | \lor I, 11 |
| 10 | Show \( \Phi \rightarrow ((\Phi \lor \Theta) \& (\Phi \lor \psi)) \) |
| 11 | \( \Phi \) | ACP |
| 12 | \( \Phi \lor (\Theta \& \psi) \) | \lor I, 11 |
| 13 | Show \( \neg \Phi \rightarrow ((\Phi \lor \Theta) \& (\Phi \lor \psi)) \) |
| 14 | \( \neg \Phi \) | ACP |
| 15 | \( \Phi \lor \Theta \) | \& E, 1 |
| 16 | Show \( \Phi \rightarrow \Theta \) |
| 17 | \( \Phi \) | ACP |
| 18 | Show \( \neg \Phi \) |
| 19 | \( \neg \Phi \) | AIP |
| 20 | \( \Phi \) | \lor R, 17 |
| 21 | \( \neg \Phi \) | \lor R, 14 |
Question: Deriving $\Phi \lor \neg \Phi$ again? Are we going to add this as a derived rule at some point?

Answer: Yes. It will be one of the miscellaneous category rules, which we'll come to below.

The arithmetic distribution law works in only one direction -- we have:
- $x(y + z) = xy + xz$
but we do not have:
- $x + yz = (x + y)(x + z)$

However, disjunction and conjunction distribute in both orientations. We've just given a rule for distributing a disjunction over a conjunction; now we will give a rule -- $\land \lor$ -- for distribution a conjunction over a disjunction. This rule has the following input-output pattern:

$$\land \lor: \begin{vmatrix}
\Phi \land (\Theta \lor \psi)
\Theta \lor \psi
\end{vmatrix}\begin{vmatrix}
(\Phi \land \Theta) \lor (\Phi \land \psi)
(\Phi \land \psi)
\end{vmatrix}$$

Again, since this is a bidirectional rule, it requires justification in both directions. Top-to-bottom, the justification is:

| (1) $\Phi \land (\Theta \lor \psi)$ | A |
| (2) Show $(\Phi \land \Theta) \lor (\Phi \land \psi)$ | &E, 1 |
| (3) $\Theta$ | &E, 1 |
| (4) $\Theta \lor \psi$ | &E, 1 |
| (5) Show $\theta \rightarrow ((\Phi \land \Theta) \lor (\Phi \land \psi))$ | ACP |
| (6) $\Theta$ | ACP |
| (7) $\Phi$ | &I, 3,6 |
| (8) $(\Phi \land \Theta) \lor (\Phi \land \psi)$ | vI, 7 |
| (9) Show $\psi \rightarrow ((\Phi \land \Theta) \lor (\Phi \land \psi))$ | ACP |
| (10) $\psi$ | ACP |
| (11) $\Phi$ | &I, 3,10 |
| (12) $(\Phi \land \Theta) \lor (\Phi \land \psi)$ | vI, 11 |
| (13) $(\Phi \land \Theta) \lor (\Phi \land \psi)$ | vE, 4,5,9 |

Bottom-to-top, the justification is:

| (1) $\Phi \land (\Theta \lor \psi)$ | A |
| (2) Show $\Phi \land (\Theta \lor \psi)$ | ACP |
| (3) Show $(\Phi \land \Theta) \rightarrow ((\Phi \land \Theta) \lor (\Phi \land \psi))$ | ACP |
| (4) $\Phi$ | &E, 4 |
| (5) $\Theta$ | &E, 4 |
| (7) $\Theta \lor \psi$ | vI, 6 |
I've given both of these distribution rules in their left-distribution format, but each of them is also valid in right-distribution format. Thus we add the following alternative forms:

\[
\begin{align*}
\lor \&: & = \Phi \lor (\Theta \lor \psi) \\
(\Theta \lor \psi) \& \phi: & = \Phi \& (\Theta \lor \psi)
\end{align*}
\]

Each of these alternative forms is easily justified using our earlier commutivity laws. Thus, for example, we can justify the new version of \(\lor \&\) in the top-to-bottom direction as follows:

1. \(\Phi \lor (\Theta \lor \psi)\) A
2. Show \(\Phi \lor (\Theta \lor \psi)\)
3. \(\Phi \lor (\Theta \lor \psi)\) \lor C, 1
4. \((\Theta \lor \psi) \lor \Phi\) \lor &, 3

Similarly, the bottom-to-top justification of the new version of \&\lor proceeds as follows:

1. \((\Theta \& \psi) \lor \Phi\) A
2. Show \((\Theta \& \psi) \lor \Phi\)
3. \((\Theta \& \psi) \lor \Phi\) & V, 1
4. \((\Theta \lor \psi) \& \Phi\) & C, 3

The justifications in the remaining two directions are essentially identical, so I won't bother giving them. It's also possible to add modified versions of the distribution rules that allow you to change the order of the smaller-scoped conjunction or disjunction, but I won't give such modifications here, since the need for them comes up only very rarely.

The idea of using older derived rules to aid in the justification of newer derived rules is one which we will come back to frequently. The use of derived rules in the justification -- as with all uses of derived rules -- is not essential (I could have justified the alternative forms of the distribution rules without using the original \(\lor \&\) and \&\lor), but it does save a considerable amount of time. By using older derived rules to justify newer derived rules, we will be able to create an ever-growing hierarchy of increasingly complex rules without needing to give justifications of derived rules growing into the thousands of lines. It's the "standing on the shoulders of giants" strategy -- we'll gradually use our older work to allow our newer work to extend farther and farther.

We've now finished the distribution rules we are going to introduce. The next set of rules are ones which convert one connective into another. For example, we know that a conditional can be expressed as a disjunction of the negation of the antecedent with the consequent, so there should be a derived rule of the following form:

\[
\begin{align*}
\phi & \rightarrow \theta \\
\rightarrow \lor: & = \neg \phi \lor \theta \\
\neg \phi & \lor \theta
\end{align*}
\]

The top-to-bottom direction of the justification is very similar to the proof we've already seen of \(\phi \lor \theta\) from \(\neg \phi \rightarrow \theta\), and proceeds as follows:

1. \(\phi \rightarrow \theta\) A
2. Show \(\neg \phi \lor \theta\)
The bottom-to-top direction of the justification is:

1. \( \neg \Phi \lor \Theta \)
2. Show \( \neg \Phi \rightarrow \Theta \)
3. \( \Phi \)
4. Show \( \Phi \rightarrow \Theta \)
5. \( \neg \Phi \)
6. Show \( \neg \Phi \rightarrow \Theta \)
7. \( \neg \Theta \)
8. \( \Phi \)
9. \( \neg \Phi \)
10. Show \( \Theta \rightarrow \Theta \)
11. \( \Theta \)
12. \( \Theta \lor E, 1,4,10 \)

The other conversion rule transforms biconditionals into conditionals. We could give a rule which directly converts a biconditional into a conjunction of conditionals, but it actually turns out to be somewhat more convenient to have a rule which simply allows us to extract a single conditional from a biconditional, like this:

\[ \Phi \leftrightarrow \Theta \quad \leftrightarrow \quad \Phi \rightarrow \Theta \]

and also like this:

\[ \Phi \leftrightarrow \Theta \quad \leftrightarrow \quad \Theta \rightarrow \Phi \]

Notice that both of these are monodirectional rules, since you cannot derive a biconditional from a single conditional. They thus each require justification only in the top-to-bottom direction (since the bottom-to-top direction is not a licensed inference). The two versions of \( \leftrightarrow \rightarrow \) have structurally identical justifications, so I will give the (quite simple) justification only of the first version:

1. \( \Phi \leftrightarrow \Theta \)
2. Show \( \Phi \rightarrow \Theta \)
3. \( \Phi \)
4. Show \( \Phi \rightarrow \Theta \)
5. \( \neg \Phi \)
6. Show \( \neg \Phi \rightarrow \Theta \)
7. \( \neg \Theta \)
8. \( \Phi \)
9. \( \neg \Phi \)
10. Show \( \Theta \rightarrow \Theta \)
11. \( \Theta \)
12. \( \Theta \lor E, 1,3 \)

The last category of rules isn't really a true category -- it's just a grab-bag of miscellaneous useful inferential moves for which we will introduce derived rules. We've already seen one entry into this category with the alternative form of conditional elimination, \( \rightarrow E^* \). We can introduce a similar alternative version of biconditional elimination, which allows us to read a biconditional "backward" with the aid of the negation of one side. The alternative version has two formulations -- one for each direction of the biconditional:

\[ \Phi \leftrightarrow \Theta \quad \leftrightarrow E^* \quad \Phi \leftrightarrow \Theta \]

Both formulations of the rule are monodirectional, and they are similar enough that I will give a justification only of the first:
We can also give an alternative version of disjunction elimination, which we will call $\lor E^*$, which can in many cases save us the trouble of going through a dilemma reasoning. The alternative version allows us to infer from a disjunction and the falsity of one disjunct to the truth of the other disjunct (since one of the two must be true). The input-output pattern is thus:

$\lor E^*$: $\phi \lor \theta$ or $\phi \lor \theta$

--------

$\theta$ $\phi$

The justifications of both forms of $\lor E^*$ are structurally identical; the first proceeds as follows:

We can also introduce a frequently-useful rule of contradiction, which allows us to avoid the work of an explicit indirect proof when we already have both halves of a contradiction on hand. The rule of contradiction, or $\bot$, has the following input-output pattern:

$\bot$:

$\phi$

$\neg \phi$

-----

$\theta$

That is, from a pair of contradictory inputs, we can derive any conclusion we want. This rule is thus a formalization of the principle of *ex falso quodlibet*. It is justified as follows:

Another convenient rule is that of the excluded middle, or $\text{EM}$. $\text{EM}$ allows us to introduce a disjunction of the form $\phi \lor \neg \phi$ at any stage in the proof. $\text{EM}$ is unique among the (non-assumption) rules we've looked at so far in requiring no inputs, and has the following input-output pattern:

$\text{EM}$: $\phi \lor \neg \phi$

We have already seen the justification for the rule of $\text{EM}$. 

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\phi \leftrightarrow \theta$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$\neg \phi$</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Show $\neg \theta$</td>
<td>AIP</td>
</tr>
<tr>
<td>4</td>
<td>$\theta$</td>
<td>AIP</td>
</tr>
<tr>
<td>5</td>
<td>$\phi$</td>
<td>$\leftrightarrow E$, 1,4</td>
</tr>
<tr>
<td>6</td>
<td>$\neg \phi$</td>
<td>R, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\phi \lor \theta$</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>$\neg \phi$</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Show $\theta$</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>Show $\phi \rightarrow \theta$</td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACP</td>
</tr>
<tr>
<td>10</td>
<td>Show $\theta \rightarrow \theta$</td>
<td>ACP</td>
</tr>
<tr>
<td>12</td>
<td>$\theta$</td>
<td>$\lor E$, 1,4,10</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\phi$</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$\neg \phi$</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Show $\theta$</td>
<td>AIP</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$\neg \theta$</td>
<td>AIP</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$\phi$</td>
<td>R, 1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$\neg \phi$</td>
<td>R, 2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\phi$</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>$\neg \phi$</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Show $\theta$</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>$\neg \theta$</td>
<td>AIP</td>
</tr>
<tr>
<td>5</td>
<td>$\phi$</td>
<td>R, 1</td>
</tr>
<tr>
<td>6</td>
<td>$\neg \phi$</td>
<td>R, 2</td>
</tr>
</tbody>
</table>
The last two derived rules we will give here provide quick methods for introducing conditionals. One rule -- which we will call \( A \rightarrow I \) -- allows us to derive a conditional from the negation of the antecedent of that conditional, and the other -- which we will call \( C \rightarrow I \) -- allows us to derive a conditional from the consequent of that conditional. We thus have:

\[
\neg \Phi \\
A \rightarrow i: \quad \text{----------} \\
\Phi \rightarrow \Theta
\]

which is justified by:

<table>
<thead>
<tr>
<th>(1) ( \neg \Phi )</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Show ( \Phi \rightarrow \Theta )</td>
<td></td>
</tr>
<tr>
<td>(3) ( \Phi )</td>
<td>ACP</td>
</tr>
<tr>
<td>(4) ( \Theta )</td>
<td>l, 1, 3</td>
</tr>
</tbody>
</table>

(note the use of the contradiction rule to shorten this justification), and also:

\[
\Theta \\
C \rightarrow i: \quad \text{----------} \\
\Phi \rightarrow \Theta
\]

which is justified by:

<table>
<thead>
<tr>
<th>(1) ( \Theta )</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Show ( \Phi \rightarrow \Theta )</td>
<td></td>
</tr>
<tr>
<td>(3) ( \Theta )</td>
<td>R, 1</td>
</tr>
</tbody>
</table>

Summary of the Derived Rules [Next]
Here is a table summarizing all of the derived proof rules I have given:

<table>
<thead>
<tr>
<th>Rearrangement Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Phi \land \Theta ) &amp; C: \quad \text{======} \quad \Theta \land \Phi</td>
</tr>
<tr>
<td>( \Phi \lor \Theta ) v C: \quad \text{======} \quad \Theta \lor \Phi</td>
</tr>
<tr>
<td>( \Phi \leftrightarrow \Theta ) &lt;-&gt; C: \quad \text{======} \quad \Theta \leftrightarrow \Phi</td>
</tr>
<tr>
<td>( \Phi \land (\Theta \land \psi) ) &amp; A: \quad \text{======} \quad (\Phi \land \Theta) \land \psi</td>
</tr>
<tr>
<td>( \Phi \lor (\Theta \lor \psi) ) v A: \quad \text{======} \quad (\Phi \lor \Theta) \lor \psi</td>
</tr>
<tr>
<td>( \Phi \leftrightarrow (\Theta \leftrightarrow \psi) ) &lt;-&gt; A: \quad \text{======} \quad (\Phi \leftrightarrow \Theta) \leftrightarrow \psi</td>
</tr>
<tr>
<td>( \Phi \rightarrow (\Theta \rightarrow \psi) ) \rightarrow: \quad \text{======} \quad (\Phi \land \Theta) \rightarrow \psi</td>
</tr>
<tr>
<td>( \Theta \rightarrow (\Phi \rightarrow \psi) ) \rightarrow Sw: \quad \text{======} \quad \Theta \rightarrow (\Phi \rightarrow \psi)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \neg(\Phi \land \Theta) ) \neg &amp;: \quad \text{======} \quad \neg(\Phi \lor \Theta)</td>
</tr>
<tr>
<td>( \neg(\Phi \lor \Theta) ) \neg v: \quad \text{======} \quad \neg(\Theta \land \neg \Theta)</td>
</tr>
<tr>
<td>( \neg(\Phi \leftrightarrow \Theta) ) \neg &lt;-&gt;: \quad \text{======} \quad \neg(\Phi \lor \neg \Theta)</td>
</tr>
<tr>
<td>( \Phi \lor (\Theta \lor \psi) ) v &amp;: \quad \text{======} \quad \Phi \lor (\Theta \land (\Theta \lor \psi) \lor \Phi)</td>
</tr>
<tr>
<td>( \Phi \land (\Theta \lor \psi) ) &amp; v: \quad \text{======} \quad \Phi \land (\Theta \lor (\Theta \lor \psi) \lor \Phi)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Connective Conversion Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Phi \rightarrow (\Theta \rightarrow \psi) ) \rightarrow &amp;: \quad \text{======} \quad (\Phi \land \Theta) \lor (\Theta \lor \psi)</td>
</tr>
<tr>
<td>( \Phi \land (\Theta \lor \psi) ) &amp; v: \quad \text{======} \quad (\Phi \land \Theta) \lor (\Theta \lor \psi)</td>
</tr>
<tr>
<td>( \Phi \lor (\Theta \lor \psi) ) v &amp;: \quad \text{======} \quad (\Phi \lor \Theta) \lor (\Theta \lor \psi)</td>
</tr>
</tbody>
</table>
This list, of course, gives only a few of the infinitely many derived rules that could be constructed. I’ve put the rules that in my experience are most useful for constructing proofs, but if you find an inferential pattern that shows up repeatedly in your proofs, you should feel free to design and justify a derived rule of your own to encode that pattern. On the flip side, I would recommend against trying to sit down and memorize all 23 of the derived rules that I’ve given here. Remember that all proofs can be done using the core rules, so knowledge of derived rules is never essential. You’re better off just working with the core rules until you notice that there are certain types of procedures that show up regularly, and then learning the derived rules which encode those procedures. That way you’ll have a clearer idea of what the derived rules accomplish for you and when they are useful.

Replacement [Next]
I’ve stressed multiple times in this discussion of the proof system the importance of using proof rules only when the connectives to which they apply appear as the main connective of the relevant input sentence. You cannot use &E on ¬(P & Q), because & is not the main connective of the input. You cannot use →E on (P → Q) ↔ R, because → is not the main connective of the input. In this section, I want to discuss an important exception to this principle.

You cannot use &E on ¬(P & Q) to get ¬P, and it’s a good thing, because ¬P doesn’t follow from ¬(P & Q). However, the way things are currently set up, you also can’t use our new derived rule &C on ¬(P & Q) to get ¬(Q & P). This is less obviously a good thing, because ¬(Q & P) does indeed follow from ¬(P & Q). Since we want to capture all of the inferences in our proof system, we want to be able to move from ¬(P & Q) to ¬(Q & P). Let’s look at how we would do this, using (but not abusing) the &C rule. The proof proceeds as follows:

| (1) ¬(P & Q) | A |
| (2) Show ¬(Q & P) | |
| (3) Q & P | AIP |
| (4) P & Q | &C, 3 |
| (5) ¬(P & Q) | R, 1 |

In this proof we do not use the rule &C directly on ¬(P & Q). Instead, we begin an indirect proof that brings a conjunction to the surface — namely, the conjunction Q & P. We then use &C on this conjunction to transform it into P & Q, which then forms a contradiction with ¬(P & Q), ending the indirect proof.

OK, so we can swap P & Q for Q & P when the conjunction is imbedded in a negation. Let’s consider another example. Suppose we want to make this inferential move:

• (P & Q) → R. ∴ (Q & P) → R

Again, we’d like to use &C to swap P & Q for Q & P, but can’t do so directly, because the conjunction is not the main connective. But we can do so indirectly:
Here again we uncover an appropriate conjunction -- Q & P -- and then use &C to transform it into P & Q, which then puts us in a position to finish the proof.

In both of the above cases, it looked initially like we wanted to use &C to move from P & Q to Q & P. We were blocked from doing so by the fact that P & Q was not at the top level of the sentence we had as input, so we had to unpack that sentence a bit. After some unpacking, we derived Q & P, and then used &C to move from it to P & Q. Notice that this is the opposite direction of &C than we were originally expecting. The important point here is this:

- If &C had been a monodirectional rule, allowing movement from P & Q to Q & P but not from Q & P to P & Q, then we wouldn't have been able to use the "backward" direction of the rule after unpacking to achieve indirectly the "forward" move from P & Q to Q & P.

Now consider another example. This time, we will try to derive the following:

\[ \neg(P \land Q) \rightarrow R, \therefore \neg(Q \land P) \rightarrow R \]

Here again we want to use &C to move from P & Q to Q & P, but this time the relevant conjunction is buried two levels deep, under a negation and a conditional. So instead we use the following proof:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>(P &amp; Q) → R</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Show (Q &amp; P) → R</td>
</tr>
<tr>
<td>3</td>
<td>Q &amp; P</td>
<td>ACP</td>
</tr>
<tr>
<td>4</td>
<td>P &amp; Q</td>
<td>&amp;C, 3</td>
</tr>
<tr>
<td>5</td>
<td>R</td>
<td>→E, 1,4</td>
</tr>
</tbody>
</table>

We use &C on line 6 once a simple conjunction has been uncovered. In this case, however, unlike the previous two we looked at, we use &C in its forward direction, to move from P & Q to Q & P.

In all three of the cases we have looked at, then, we have eventually been able to use &C as part of the process of making the transition from P & Q to Q & P, even when P & Q is buried under other connectives. In some cases, we have used &C "backward", to move from Q & P to P & Q. In other cases, we have used it forward, to move from P & Q to Q & P. It has thus been crucial that &C be a bidirectional rule, allowing us to move in either of these directions. Had &C been monodirectional allowing only the forward direction, we would have been unable to swap P & Q for Q & P when P & Q was in the immediate scope of a main connective of \( \rightarrow \) or \( \rightarrow \). Had &C been monodirectional allowing only the backward direction, we would have been unable to swap P & Q for Q & P when P & Q was in the immediate scope of a negation in the immediate scope of a main connective of \( \rightarrow \) (or when P & Q occurred on its own). But with a bidirectional &C, all of these cases can be handled.

On the basis of the examples we’ve looked at so far, I am going to venture a bold hypothesis:

- If \( \Sigma \) is a sentence containing a subsentence \( \Phi \) to which a bidirectional rule R applies to produce output \( \psi \), it is always possible to create a proof that derives \( \Sigma[\psi/\Phi] \) from \( \Sigma \).

For example, let \( \Sigma \) be \( \neg(P \land Q) \rightarrow R \). \( \Sigma \) then contains \( \Phi = P \land Q \), to which the bidirectional rule of &C applies to produce output \( \psi = Q \land P \). The bold hypothesis then holds that it is possible to create a proof which derives \( \neg(Q \land P) \rightarrow R \) from \( \neg(P \land Q) \rightarrow R \). In this particular case, we have already seen that the bold hypothesis is correct.

I want to emphasize right away that the evidence we’ve looked at so far is by no means adequate to establish the truth of the bold hypothesis. The hypothesis is a claim about an infinite number of cases, and we’ve looked at only three so far. A full proof of the hypothesis would require using techniques of mathematical induction that we haven’t come to yet. Nevertheless, the hypothesis is in fact true. (You’ll just have to take my word for it now -- we’ll come back to this issue later and provide a proof of it.)

On the basis of this bold hypothesis, we are going to adopt a new principle for performing proofs. This is the principle of Replacement:
Replacement: A bidirectional proof rule can be used any time its required input shows up anywhere, not just when its required input is at the top level of logical structure. Therefore, if $\Sigma$ is a sentence containing a subsentence $\Phi$ to which a bidirectional rule $R$ applies to produce output $\psi$, we can use rule $R$ to move directly from $\Sigma$ to $\Sigma[\psi/\Phi]$.

The principle of Replacement thus tells me that in each of the three cases, I can simply directly apply $\&C$ to $P \& Q$ without bothering to "uncover" it. Thus the following three proofs are all legal:

1. $\neg(P \& Q)$
2. Show $\neg(Q \& P)$
3. $\neg(Q \& P)$

1. $(P \& Q) \rightarrow R$
2. Show $(Q \& P) \rightarrow R$
3. $(Q \& P) \rightarrow R$

1. $\neg(P \& Q) \rightarrow R$
2. Show $\neg(Q \& P) \rightarrow R$
3. $\neg(Q \& P) \rightarrow R$

Replacement can be a tremendous time-saver, allowing you to use proof rules directly on deeply imbedded sentences without doing any unpacking of those sentences.

It is very important, however, that you keep in mind that Replacement only allows you to use bidirectional rules in this manner. Monodirectional rules can still be applied only when their inputs appear as isolated sentences in the proof, not as imbedded components of other sentences. Among the core proof rules, all but $\neg\neg$ are monodirectional, so only $\neg\neg$ can be used via Replacement in imbedded contexts (you can eliminate or introduce a double negative anywhere you want). Many, but not all, of the derived rules I've introduce are bidirectional. All of the bidirectional rules can, via Replacement, be used on deeply imbedded inputs; none of the monodirectional ones can be.

Let's now look at an example to show how Replacement can save time. We'll prove the following result:

1. $(P \leftrightarrow Q) \leftrightarrow (R \leftrightarrow S)$
2. $(P \leftrightarrow R) \leftrightarrow (Q \leftrightarrow S)$

The proof proceeds as follows:

1. $(P \leftrightarrow Q) \leftrightarrow (R \leftrightarrow S)$
2. Show $(P \leftrightarrow R) \leftrightarrow (Q \leftrightarrow S)$
3. $P \leftrightarrow (Q \leftrightarrow (R \leftrightarrow S))$
4. Show $P \leftrightarrow ((Q \leftrightarrow R) \leftrightarrow S)$
   a. $P \leftrightarrow A$, 1
5. $P \leftrightarrow ((R \leftrightarrow Q) \leftrightarrow S)$
6. $P \leftrightarrow (R \leftrightarrow (Q \leftrightarrow S))$
7. $P \leftrightarrow A$, 5
8. $(P \leftrightarrow R) \leftrightarrow (Q \leftrightarrow S)$

In this proof, the principle of Replacement is appealed to in lines 4, 5, and 6. In each of these lines, a bidirectional rule (either $\leftrightarrow A$ or $\leftrightarrow C$) is used on a sentence which is imbedded in some larger context. A considerable amount of effort is saved by using Replacement in this way. Consider, for example, the move from line 3 to line 4. Without Replacement, we would have had to unpack the lower biconditional in order to use $\leftrightarrow A$. The procedure would have looked like this:

...
The use of Replacement thus saves eight lines in that step alone. The savings are the same in the move from line 5 to line 6, and even greater in the move from line 4 to line 5 (since two levels of biconditional have to be unpacked here). Our seven-line proof would unpack to about 35 lines if we didn't have Replacement available. (And, of course, if we didn't have the derived rules, it would be much worse. Each use of ↔C saves six lines, and each use of ↔A saves an impressive 83 lines, and the number of applications of these rules doubles for singly-imbedded Replacement applications and quadruples for doubly-imbedded Replacement applications, so the proof done with the core rules would run an unpleasant 391 lines.)

Some Examples Using Derived Rules

I am now going to give five examples of proofs using the derived proof rules, and Replacement with derived proof rules where appropriate, to help illustrate how they shorten proofs. In the first example, I will give both the proof with derived rules and the proof without derived rules, to make the difference as clear as possible. In the remaining four examples, I'll just give a quick indication of the length of the proof done without derived rules.

For our first example, we will prove the following result:

\[(P \leftrightarrow Q) \leftrightarrow R, \neg(P \leftrightarrow \neg R), (Q \lor \neg S) \rightarrow T, \therefore T\]

The proof proceeds as follows:

1. \((P \leftrightarrow Q) \leftrightarrow R\) A
2. \(\neg(P \leftrightarrow \neg R)\) A
3. \((Q \lor \neg S) \rightarrow T\) A
4. Show T
   5. \(T\) AIP
   6. \(\neg(Q \lor \neg S)\) \(\rightarrow E^*, 3,5\)
   7. \(\neg Q \& \neg\neg S\) \(\rightarrow \& E, 6\)
   8. \(\neg Q\) \(\neg \& E, 7\)
   9. \(\neg P \leftrightarrow \neg \neg R\) \(\leftrightarrow E*, 2\)
10. Show \(\neg R\)
   11. \(R\) AIP
   12. \(P \leftrightarrow Q\) \(\leftrightarrow E, 1,11\)
   13. \(\neg P\) \(\leftrightarrow E, 12,8\)
   14. \(\neg R\) \(\leftrightarrow E, 9,13\)
15. \(\neg P\) \(\leftrightarrow E, 9,10\)
16. \(P \leftrightarrow Q\) \(\leftrightarrow E^*, 1,10\)
17. \(P \leftrightarrow Q\) \(\leftrightarrow E, 16\)
18. \(Q\) \(\leftrightarrow E, 17,15\)

Without use of derived rules, we have to unpack one occurrence of \(\rightarrow E^*\), one occurrence of \(\neg \&\), two occurrences of \(\neg \leftrightarrow\), and two occurrences of \(\leftrightarrow E^*\). The resulting rather monstrous proof is:

1. \((P \leftrightarrow Q) \leftrightarrow R\) A
2. \(\neg(P \leftrightarrow \neg R)\) A
3. \((Q \lor \neg S) \rightarrow T\) A
4. Show T
   5. \(T\) AIP
   6. \(\neg(Q \lor \neg S)\) \(\rightarrow E, 3,7\)
   7. \(\neg T\) \(R, 5\)
10. Show \(\neg Q\)
   11. \(Q\) AIP
   12. \(Q \lor \neg S\) \(\lor I, 11\)
   13. \(\neg(Q \lor \neg S)\) \(R, 6\)
14. Show \(P \rightarrow \neg R\)
   15. \(P\) ACP
   16. Show \(\neg R\)
   17. \(R\) AIP
   18. Show \(P \rightarrow \neg R\)
   19. \(P\) ACP
(20) Show \( \neg R \)

(21) R AIP

(22) P R, 19

(23) \( \neg P \) R, 15

(24) Show \( \neg R \rightarrow P \)

(25) \( \neg R \) ACP

(26) Show P

(27) \( \neg P \) AIP

(28) R R, 21

(29) \( \neg R \) R, 25

(30) P \( \leftrightarrow \neg R \) \( \leftrightarrow \) 18, 24

(31) \( \neg (P \leftrightarrow \neg R) \) R, 2

(32) Show \( \neg R \rightarrow \neg P \)

(33) \( \neg R \) ACP

(34) Show \( \neg P \)

(35) P AIP

(36) Show P \( \rightarrow \neg R \)

(37) P ACP

(38) \( \neg R \) R, 33

(39) Show \( \neg R \rightarrow P \)

(40) \( \neg R \) ACP

(41) P R, 35

(42) P \( \leftrightarrow \neg R \) \( \leftrightarrow \) 36, 39

(43) \( \neg (P \leftrightarrow \neg R) \) R, 2

(44) \( \neg P \leftrightarrow \neg R \) \( \leftrightarrow \) 14, 32

(45) Show \( \neg R \)

(46) R AIP

(47) P \( \leftrightarrow Q \) \( \leftrightarrow \) E, 1, 46

(48) Show \( \neg P \)

(49) P AIP

(50) Q \( \leftrightarrow \) E, 47, 49

(51) \( \neg Q \) R, 10

(52) \( \neg R \) \( \leftrightarrow \) E, 44, 48

(53) \( \neg P \) \( \leftrightarrow \) E, 44, 45

(54) Show \( \neg (P \leftrightarrow Q) \)

(55) P \( \leftrightarrow Q \) AIP

(56) R \( \leftrightarrow \) E, 1, 55

(57) \( \neg R \) R, 45

(58) Show \( \neg P \rightarrow Q \)

(59) \( \neg P \) ACP

(60) Show Q

(61) \( \neg Q \) AIP

(62) Show P \( \rightarrow Q \)

(63) P ACP

(64) Show Q

(65) \( \neg Q \) AIP

(66) P R, 63

(67) \( \neg P \) R, 59

(68) Show Q \( \rightarrow P \)

(69) Q ACP

(70) Show P

(71) \( \neg P \) AIP

(72) Q R, 69

(73) \( \neg Q \) R, 61

(74) P \( \leftrightarrow Q \) \( \leftrightarrow \) 62, 68

(75) \( \neg (P \leftrightarrow Q) \) R, 54

(76) Show Q \( \rightarrow \neg P \)
This proof should be enough to dispel any remaining skepticism about the utility of the derived proof rules.

For our second example, we will prove the following result:

• \((P & Q) \iff R, \neg(R \lor \neg P), \therefore \neg Q\)

The proof proceeds as follows:

(1) \((P & Q) \iff R\) A
(2) \(\neg(R \lor \neg P)\) A
(3) Show \(\neg Q\)
(4) \(\neg R \& \neg \neg P\) \(\neg \lor, 2\)
(5) \(\neg R\) \&E, 4
(6) \(\neg(P & Q)\) \(\to \lor^*, 1,5\)
(7) \(\neg P \lor \neg Q\) \(\neg \&^*, 6\)
(8) \(\neg P\) \&E, 4
(9) \(\neg Q\) \(\lor \lor^*, 7,8\)

Note that the use of the indirect proof rules here allows us to avoid having any subproof structure (since all the subproof structure from the underlying proof with the core rules is absorbed into the derived rules). Without use of derived rules, this proof would run 42 lines, so the derived rules strip off almost 3/4 of the length of the original proof.

For our third example, we will prove the following result:

• \(\therefore (P \iff Q) \lor (P \iff R) \lor (Q \iff R)\)

This result tells us that if we have three sentence letters, some two of them must have the same truth value. The proof proceeds as follows:

(1) Show \((P \iff Q) \lor (P \iff R) \lor (Q \iff R)\) A
(2) \(\neg((P \iff Q) \lor (P \iff R) \lor (Q \iff R))\) A
(3) \((P \iff Q) \& \neg(P \iff R) \& \neg(Q \iff R)\) \(\neg \lor, 2\)
(4) \((P \iff Q)\) \&E, 3
(5) \(P \iff \neg Q\) \(\iff \iff, 4\)
(6) \((P \iff R)\) \&E, 3
(7) \(P \iff \neg R\) \(\iff \iff, 6\)
(8) \((Q \iff R)\) \&E, 3
(9) \(Q \iff \neg R\) \(\iff \iff, 8\)
(10) Show \(Q\)
(11) \(\neg Q\) A
(12) \(P\) \(\iff \lor, 5,11\)
(13) \(\neg R\) \(\iff \lor, 7,12\)
(14) \(Q\) \(\iff \lor, 9,13\)
(15) Show \(\neg Q\)
(16) \(Q\) A
(17) \(\neg R\) \(\iff \lor, 9,16\)
(18) \(P\) \(\iff \lor, 7,17\)
(19) \(\neg Q\) \(\iff \lor, 5,18\)
The use of \(\neg\) and \(\leftrightarrow\) here saves a total of 103 lines in the proof, and thus eliminates over 75% of the proof.

For our fourth example, we will provide the justification (which we omitted above) of the \(\leftrightarrow\) rule in the top-to-bottom direction. We thus need to prove:

\[
\Phi \leftrightarrow (\Theta \leftrightarrow \psi), \therefore (\Phi \leftrightarrow \Theta) \leftrightarrow \psi
\]

The proof proceeds as follows:

\[
\begin{align*}
1. & \quad \Phi \leftrightarrow (\Theta \leftrightarrow \psi) \\
2. & \quad \text{Show } (\Phi \leftrightarrow \Theta) \leftrightarrow \psi \\
3. & \quad \text{Show } (\Phi \leftrightarrow \Theta) \rightarrow \psi \\
4. & \quad \Phi \leftrightarrow \Theta \\
5. & \quad \phi \rightarrow \psi \\
6. & \quad \text{Show } \phi \rightarrow \psi \\
7. & \quad \phi \\
8. & \quad \Theta \leftrightarrow \psi \\
9. & \quad \Theta \\
10. & \quad \psi \\
11. & \quad \text{Show } \neg \phi \rightarrow \psi \\
12. & \quad \neg \phi \\
13. & \quad \neg(\Theta \leftrightarrow \psi) \\
14. & \quad \Theta \rightarrow \neg \psi \\
15. & \quad \Theta \\
16. & \quad \neg \psi \\
17. & \quad \psi \\
18. & \quad \psi \\
19. & \quad \text{Show } \psi \rightarrow (\Phi \leftrightarrow \Theta) \\
20. & \quad \psi \\
21. & \quad \text{Show } \phi \rightarrow \Theta \\
22. & \quad \phi \\
23. & \quad \Theta \leftrightarrow \psi \\
24. & \quad \Theta \\
25. & \quad \text{Show } \Theta \rightarrow \phi \\
26. & \quad \Theta \\
27. & \quad \psi \rightarrow \Theta \\
28. & \quad \psi \\
29. & \quad \Theta \leftrightarrow \psi \\
30. & \quad \Theta \\
31. & \quad \phi \leftrightarrow \Theta \\
32. & \quad \phi \leftrightarrow \psi \\
\end{align*}
\]

This proof is none too brief even with the derived rules, but without the two use of \(C \rightarrow I\), one use of \(EM\), three uses of \(\leftrightarrow E^*\), and one use of \(\neg \leftrightarrow\), it would have been 53 lines longer.

For our fifth and final exam, we will prove the following result:

\[
\neg((P \lor Q) \leftrightarrow ((P \rightarrow Q) \rightarrow Q)), \therefore R
\]

The proof proceeds as follows:

\[
\begin{align*}
1. & \quad \neg((P \lor Q) \leftrightarrow ((P \rightarrow Q) \rightarrow Q)) \\
2. & \quad \text{Show } R \\
3. & \quad \text{Show } (P \lor Q) \leftrightarrow ((P \rightarrow Q) \rightarrow Q) \\
4. & \quad P \lor Q \\
5. & \quad Q \lor \neg Q \\
6. & \quad (P \lor Q) \lor (Q \lor \neg Q) \\
7. & \quad (P \lor Q) \lor (\neg Q \lor Q) \\
8. & \quad (P \lor \neg Q) \lor Q \\
9. & \quad (P \rightarrow Q) \lor Q \\
10. & \quad (P \rightarrow Q) \rightarrow Q \\
11. & \quad ((P \rightarrow Q) \rightarrow Q) \rightarrow (P \lor Q) \\
\end{align*}
\]
In this proof, we use EM once, $\lor C$ once, I once, $\lor \&$ twice, $\rightarrow \lor$ twice, and $\leftrightarrow$ twice, for a total savings of 98 lines.

Why Do I Have to do All These Meaningless Proofs With Ps and Qs? [Next]

At this point, you should (if you've been following everything that's been done so far and been practicing with the proof system) have a very thorough mastery of proofs in sentential logic. It's now time to step back and spend some time addressing a question which may have been nagging at you for some time now: why do we want to be able to construct all these proofs? I am not going to try to address the general question of why we want a proof system at all, since I've already given what answer I can to that question earlier. Instead, I am going to address the feeling (which seems to be quite a common one) that there is something particularly meaningless about the kind of proofs we are doing.

It's certainly easy to see how this feeling could arise. It's one thing to produce a proof justifying an argument like:

- A fetus is a person if and only if a convicted murderer is.
- Abortion is the killing of a fetus and capital punishment is the killing of a convicted murderer.
- Both abortion and capital punishment are morally wrong if and only if they involve the killing of a person.

- $\therefore$ If abortion is morally permissible, so is capital punishment.

or like:

- Given any line and a point not on that line, there is exactly one line parallel to the given line passing through the given point.
- Given triangle ABC, C is not on the line AB.
- Vertical angles are congruent.

- $\therefore$ The sum of the angles in a triangle is $180^\circ$.

In both of these cases, it's clear that we are producing a proof of real claims of real interest. However, when you are asked instead to produce a proof of a claim like:

- $\neg (\neg P \& Q) \rightarrow (P \leftrightarrow Q)$

and then spend an hour struggling through the production of a thirty-odd line proof, it can be quite natural to wonder why it's worth putting such effort into rearranging a bunch of meaningless letters.

What I want to do, then, is to explain why you really are doing something worth doing as you struggle through these apparently meaningless proofs. I will give five justifying reasons. I can't promise that these reasons will completely dispel a sense of dissatisfaction with sentential proofs, but hopefully they will give at least some idea of why this is worth learning.

Reason #1: Proofs in sentential logic are just abstract versions of particular proofs in English.

Since we can think of each sentence letter in the sentential logic as standing for some sentence of English, these apparently meaningless proofs can all be thought of as particular, concrete proofs in English raised to a level of abstraction. If we adopt the following translation key:

- P: A fetus is a person.
- Q: A convicted murderer is a person.
- R: Abortion is the killing of a fetus
- S: Capital punishment is the killing of a convicted murderer.
- T: Abortion is morally permissible.
- U: Capital punishment is morally permissible.
- V: Abortion is the killing of a person.
- W: Capital punishment is the killing of a person.

Then the argument we gave above translates into the formal logic as:

- $P \leftrightarrow (Q \& R) \& (\neg T \leftrightarrow V) \& (\neg U \leftrightarrow W) \& (P \& Q) \leftrightarrow V \& (Q \& S) \leftrightarrow W \therefore T \rightarrow U$

(Note that I've added here two auxilliary assumptions $(P \& R) \leftrightarrow V$ and $(Q \& S) \leftrightarrow W$, which serve to link (in the first case, analogously in the second case) the claims that a fetus is a person and abortion is the killing of a fetus with the claim that abortion is the killing of a person. It's a good exercise to think about why these extra assumptions are needed here.) This formal argument can then be given the following proof:

| (1) $P \leftrightarrow Q$ | A |
| (2) $R \& S$ | A |
We could put this proof into English as follows (omitting the statement of the assumptions):

Suppose that abortion is morally permissible. Since abortion is morally wrong if and only if it is the killing of a person, it follows from our supposition that abortion is not the killing of a person. But abortion is the killing of a person if and only if abortion is the killing of a fetus and the fetus is a person. Since abortion is not the killing of a person, those can't both be right, so either abortion is not the killing of a fetus, or the fetus is not a person. But abortion is the killing of a fetus, so it follows that the fetus is not a person. However, the fetus is a person if and only if a convicted murderer is, so it follows that a convicted murderer also is not a person. Since it's not true that a convicted murderer is a person, it's also not true that both the convicted murderer is a person and capital punishment is the killing of a convicted murderer. Since capital punishment's being the killing of a person requires both of these conditions to hold, and since they don't both hold, it follows that capital punishment is not the killing of a person. But capital punishment is wrong if and only if it is the killing of a person -- since it is not, it follows that capital punishment is not wrong, and is morally permissible.

You should by now be able to match the steps in this informal proof in English with the steps in the formal proof above, and see how the structure of the two correspond.

The hope here is that seeing how the above formal proof is just an abstracted version of a particular proof in English will help bring that proof to life, and make it seem less pointless. The same idea can be pursued in reverse when you start with a formal proof. Suppose you've been asked to prove the following claim:

\[ P \rightarrow \neg(Q \land 

(Remember, here you see only one side of each of the four cards.) I now give the following rule:

Every card with a 2 on one side must have a C on the other side.

The question now is: which cards must you turn over in order to determine whether my rule is being followed?

Almost everyone sees, correctly, that the first card -- the one showing a 2 -- must be turned over. If this card does not have a C on the other side, then the rule is not followed, so to check obedience to the rule we need to check that there is a C on the other side of the card. Many people decide incorrectly that the second card -- the one showing a C -- must also be turned over. This is incorrect because this card is irrelevant to adherence to the rule. If the other side of this card is a 2, then it is in conformity with the rule (having a 2 on one side and a C on the other), and if the other side is not a 2, it is still in conformity with the rule (since the rule places no restrictions on what can show up on the flip side of numbers other than 2). Relatively few people see that it is also necessary to turn over the third card -- the one showing an F. If this card has a 2 on the other side, then the rule has not been obeyed, since this would then be a card with a 2 on one side and without a C on the other side.

The correct answer, then, is that the first and third cards need to be turned over. This problem is known in the literature on rationality as the Wason Selection Task, and repeated trials have shown that only about 25% of typical respondents will give the correct answer. However, the results are quite different with an only slightly modified version of the task. Suppose now that I have a collection of checks, each of which has an amount on the front and then either an endorsement or no endorsement on the back. I place four checks in front of you (some right-side up, some upside down), revealing the following:

$500 (endorsement)  $1200 (no endorsement)

I now give the following rule:

Every check for more than $1000 must have an endorsement on the other side.

The question now is: which checks must you turn over in order to determine whether my rule is being followed?

In this case, the answer is that the third and fourth checks need to be turned over. The third check needs to be turned over, because if it is unendorsed, it violates the rule. The fourth check needs to be turned over, because if it is for an amount over $1000, it violates the rule. In this kind of example, about 75% of typical respondents give the correct answer.

The quite marked difference between people's success rates at the two tasks is interesting, because formally the two tasks are perfectly equivalent. In each case, we are being asked to check the truth of a certain conditional of the form \( P \rightarrow Q \). In the first case, we have:

- \( P = \) There is a 2 on one side of the card.
- \( Q = \) There is a C on the other side of the card.

In the second case, we have:

- \( P = \) There is an amount of over $1000 on one side of the check.
- \( Q = \) There is an endorsement on the other side of the check.

We want to see if there are any cases (cards or checks) for which the conditional comes out false. The only way for a conditional to be false is for the antecedent to be true and the consequent to be false -- that is, for \( P \) & \( \neg Q \) to hold. Of course, for any given card or check, we only get to see one side initially. There are thus four possible cases we can see -- \( P, \neg P, Q, \neg Q \). If we see a \( P \) or a \( \neg Q \), we need to check the other side, because each of these cases could be appropriately completed to falsify the conditional. However, if we see a \( \neg P \) or a \( Q \), then there is no need to check the other side, because no matter what is on the other side, it could not combine with the side we've already seen to form \( P \) & \( \neg Q \) and falsify the conditional.

So in the first case, we need only check cards which meet one of these two conditions:

- Showing a 2 on one side (making \( P \) true)
- Showing a letter other than C on one side (making \( \neg Q \) true)

In the second case, we need only check checks which meet one of these two conditions:

- Showing an amount of over $1000 on one side (making \( P \) true)
- Showing an unendorsed back side (making \( \neg Q \) true)

The temptation which traps many people to insist that the card showing a C needs to be checked can now easily be seen to be misleading. A card showing a C is one which makes \( Q \) true, and making \( Q \) true cannot possibly lead to making \( P \rightarrow Q \) false, since the truth of the consequent guarantees the truth of the conditional.
If the two tasks are structurally equivalent, why is there such a disparity in people’s success rates at performing them? The most likely explanation seems to be that the second task is designed as a concrete situation of a sort that people are used to dealing with, while the first situation is abstract, contextless, and unusual. People are used to dealing with various bureaucratic requirements that various documents be signed in various ways in various situations — they have performed similar tasks frequently before, so the present task is largely a matter of bringing to bear skills that they’ve already mastered, adapted slightly for the peculiarities of the particular situation. The first task, however, is abstract, contextless, and pointless. There’s no apparent reason for the rule that cards with a 2 on one side have a C on the other side, and this lack of motivation makes it harder for us to engage cognitively with the exercise.

The lesson of the Wason Selection Task, then, is that people are better at concrete reasoning about situations they are accustomed to than they are at abstract and contextless reasoning. The kinds of proofs we have been doing here, unfortunately, are practically a caricature of abstract and contextless reasoning. It is in part for this reason that people often find the proof system hard to work with. Knowing this, we can then see that the problem can be ameliorated by giving the sentence letters in a proof an interpretation, so that the proof ceases to be so abstract.

While the abstractness of the proof system is a defect insofar as it places it in the cognitive blind spot identified by the Wason Selection Task, there is also a good side to that abstractness. For one thing, abstraction is valuable in allowing us to see patterns that might have escaped us had we remained in the concrete. Compare, for example, the following two informal arguments:

**Argument #1:** Only if there are large spending increases for national defense will both Republicans and conservative Democrats support the president’s budget proposal, but such increases are impossible in the current economic climate. Since the Republicans will definitely support the president’s proposal in any case, the conservative Democrats will not break rank with their party to back the president.

**Argument #2:** The number of neutrons in this isotope is below 200. We know this because if the isotope had at least 200 neutrons and at most 95 protons, it would be stable. However, it is an unstable sample of uranium, which has 92 protons.

A little examination will show that each of these two arguments has the same underlying structure:

- \((P \land Q) \rightarrow R\).
- \(-R \land P \land Q\).
- \(\therefore \neg P\)

Without the level of abstraction that the formal logic provides, we might not have noticed that the two arguments had the same structure, and hence that they must stand or fall together (stand, in this case). For another thing, the very fact that people are bad at abstract reasoning provides a reason to do more of it, so that we can shore up our weaknesses.

**Reason #2: Sentential proofs capture the structure of real mathematical proofs**

Consider the following proof, taken from a text on topology:

**Theorem 9:** Let \(K\) be an arcwise connected graph. Let \(n\) be the maximum number of open 1-simplices which can be removed from \(K\) without disconnecting the space. \((n\) is the number of “basic” circuits in \(K\)). Then \(n = 1 - \chi(K)\).

**Proof:** If \(K\) is a tree, then \(n = 0\), and Theorem 9 applies. If \(K\) is not a tree, let \((s_1)\) be an open 1-simplex such that \([K] – (s_1)\) is connected. If \(K – (s_1)\) is a tree, stop. Otherwise, let \((s_2)\) be an open 1-simplex such that \([K] – (s_1) \cup (s_2)\) is connected. Continue. Since there are only finitely many 1-simplices in \(K\), the process must stop; that is, for some \(n, K – \{(s_1), (s_2), ..., (s_n)\}\) is a tree \(T\). Then:

\[
\chi(K) = \chi(T) – n = 1 – n;
\]

that is,

\[
n = 1 - \chi(K).
\]

In all likelihood, not only the workings of this proof but even what the original theorem is saying will be opaque to most readers. However, even without being able to follow these matters, certain features of the proof should be comprehensible and, in fact, quite familiar. The proof opens by offering a choice between two options: \(K\) is a tree, or \(K\) is not a tree. This is an appeal to the rule of **EM**. It then uses this disjunction (in the same way that we’ve done multiple times before) to trigger a bit of dilemma reasoning. One branch of the dilemma (\(K\) is a tree) terminates immediately. Why it terminates may be mysterious without knowing the relevant math, but we can see that that branch is treated as complete. In the other branch, **EM** is invoked again, this time to offer a choice between \(K – (s_1)\) being a tree, and \(K – (s_2)\) being a tree. Again the disjunction produced through **EM** triggers a dilemma reasoning. Again one branch – \(K – (s_1)\) is a tree – terminates immediately, while the other leads to yet another use of **EM** (this time producing \(K – \{(s_1), (s_2)\}\) is a tree or \(K – \{(s_1), (s_2)\}\) is not a tree), which leads to another dilemma reasoning, and so on. As a result, an extremely elaborate construction of nested dilemma reasonings is created. It’s then claimed that this process can’t go on infinitely long, so all of the branches of the many dilemmas must all eventually terminate, thereby successfully concluding the proof. Without knowing any of the math involved, we can see that the proof of this theorem is a living example of the proof techniques we’ve been developing here.
Theorem 2.11.1: If G is a finite group, then ca = o(G)/o(N(a)); in other words, the number of elements conjugate to a in G is the index of the normalizer of a in G.

Proof: Suppose that x, y ∈ G are in the same right coset of N(a) in G. Thus  y = nx, where n ∈ N(a), and so na = an. Therefore, since y⁻¹y = (nx)⁻¹ = x⁻¹n⁻¹, y⁻¹ay = x⁻¹n⁻¹anx = x⁻¹n⁻¹nax = x⁻¹ax, whence y and x result in the same conjugate of a.

If, on the other hand, x and y are in different right cosets of N(a) in G we claim that x⁻¹ax ≠ y⁻¹ay. Were this not the case, from x⁻¹ax = y⁻¹ay we would deduce that yx⁻¹a = ayx⁻¹; this in turn would imply that yx⁻¹ ∈ N(a). However, this declares x and y to be in the same right coset of N(a) in G, contradicting the fact that they are in different cosets.

Again, the mathematical details are probably obscure. But there are structural elements of this proof that should look familiar. An initial disjunction is proposed, in keeping with EM – either x and y are in the same right coset of N(a) (whatever that means), or they are in different cosets. A dilemma reasoning is then invoked. The first branch of the dilemma in essence traces along a sequence of conditionals using →E to create the series of equalities which terminates that branch. The second branch invokes indirect proof, assuming the negation of what is to be shown – that is, assuming x⁻¹ax = y⁻¹ay – and arguing from there (again by a sequence of applications of →E) to reach a contradiction.

In neither of these cases does the proof system of sentential logic allow us to follow all of the details of the proof (see Reason #4 below for more on this point), but it does allow us to recognize some of the large features of the proof. Sentential proof rules are an important part of the language of mathematics – the language of mathematics is an important one, and sentential proofs then inherit importance from their involvement in that language.

Reason #3: Proof in sentential logic are just the beginning of the larger program

In giving the previous reason, we saw that the proofs we've been doing in sentential logic can be detected inside the workings of real mathematical proofs. However, there is still much in each of the two proofs I quoted that we cannot capture using the proof system we've developed. In the first proof, we can "see", with the goggles we've designed, the iterative dilemma reasoning skeleton, but we cannot capture, for example, the reasoning that says that there can be at most finitely many 1-simplicies extracted from K, or the reasoning that moves from χ(K) = χ(T) – n = 1 – n to n = 1 - χ(K). In the second argument, we can see the dilemma reasoning structure, and the indirect proof contained in the second branch of the dilemma, but we cannot capture, for example, the move from x and y being in the same right coset of N(a) to the claim that y = nx, or the series of arithmetic transitions in the first branch of the dilemma.

We can't capture these aspects of the mathematical proofs because we are only at the very beginning of the process of developing logic and a proof system. While we do have a fully developed proof system for sentential logic, that logic itself is a rather crude tool, and allows us to capture only the roughest outlines of real proofs. One of the reasons, then, that proofs in sentential logic create a feeling of pointlessness is because we are still so limited in what we can do. It is as if we had set out to become master artists -- so far, we've had intensive lessons on how to create a canvas, going through excruciating detail on all the methods of weaving the fibers to make canvases of various styles. Now, having a canvas is a rather crucial part of creating a masterpiece of art, and if the making of the canvas is in fact just the first step in going on the paint the Mona Lisa, then it's not to hard to see how the canvases-making can be a rewarding part of the process. But if you're stuck in the apprentice's lab, spending month after month doing nothing but making canvas after canvas, then of course the process is going to get old quickly, and the canvases will come to seem pointless drudgery. However, you can't make the masterpiece without the canvas -- learning the preliminary steps may be unpleasant, but it's an essential step toward eventually reaching the desired goal.

The sentential proof system is only the skeleton of the Frankenstein's monster we are in the process of building. A skeleton is, of course, a crucial part of a good monster, but it won't be until we add some muscles and skin, and that brain that Igor is hunting down for us, that we will fully realize the value of the osteological time investment we've been making here.

Reason #4: The proof system for sentential logic is needed so that we can prove claims about proving claims

In the previous three reasons, I've tried to give some justification for the proof system involving Ps and Qs as a method (albeit only a first step in a more complete method) of capturing abstract and highly general features of real proofs we do in real life. For my last reason, I want to make a still-further step of abstraction. So far we have considered how the sentential logic proof system can be useful by having its abstract proofs capture the structure of a category of concrete proofs we are interested in. Now I want to consider how the proof system allows us to go beyond performing particular proofs (whether concrete or abstract) to the more interest task of proving claims about proving claims. Now that we've got a proof system which is defined with full precision, we have perfectly precise standards for what counts and what doesn't
count as a proof. Given those claims, we can now ask, and answer, very general questions about the task of producing proofs. We can address questions such as:
- Is there an algorithmic method for producing proofs?
- Is there a way of determining constraints on how long a proof is needed to establish a claim?
- Can we prove everything in the proof system we ought to be able to prove?
- What sorts of conditions on interpretations can we impose by holding true various claims in the sentential logic
- and so on

What we are in a position to do, then, is to move into an area of logic called metatheory. In metatheory, our goal is not to produce proofs in the proof system, but to consider the proof system, and the logical language for which it is a proof system, as a mathematical entity, and then to investigate the theoretical properties of that mathematical entity. Here’s a very simple example of a metatheoretic result:

**Deduction Theorem:** Θ is derivable from Φ if and only if Φ → Θ is a theorem (that is, is derivable from no assumptions).

**Proof:** Since the Deduction Theorem is a biconditional, it must be proved in both directions. We’ll start with the left-to-right direction. Assume, then, that Θ is derivable from Φ. That is, there is some proof Π which has the following structure:

| (1) Φ | A |
| (2) Show Θ | |
| ... |

which successfully derives Θ from the single assumption Φ. We now want to show that Φ → Θ can be derived from no assumptions. To do so, start with the following proof fragment (which we will call Ω):

| (1) Show Φ → Θ |
| (2) Φ | ACP |

Now modify Π by removing its first line and increasing all subsequent line numbers by 1 (including citations of these line numbers as inputs to proof rules, creating a new proof fragment Π*. Now paste together Ω and Π*. The result will be a proof of Φ → Θ, since Ω will (on line 2) provide the assumption of Φ that Π* needs, and Π* will then take that assumption and use it to derive the Θ that Ω needs to cancel the Show of Φ → Θ.

In the right-to-left direction, assume we have some proof Π of Φ → Θ from no assumptions. We now want to show that we can derive Θ from Φ. We construct a proof by starting with this proof fragment:

| (1) Φ | A |
| (2) Show Θ | |

We then modify Π by increasing all line numbers by 2, and paste it onto the above fragment, to give:

| (1) Φ | A |
| (2) Show Θ | |
| (3) Show Φ → Θ | |
| ... [the rest of Π] |

Once we’ve derived Φ → Θ, a final →E finishes the proof of Θ from Φ:

| (1) Φ | |
| (2) Show Θ | |
| (3) Show Φ → Θ | |
| ... [the rest of Π] |
| (n) Θ | →E, 1,3 |

While this is hardly the deepest result that metatheoretical investigation will give us, it still reveals an interesting connection between the conditional and the notion of provability, and thus helps justify our informal practice as reading → as implies, when we do so cautiously. Notice, for example, that we could not have proved a similar result showing that Θ
was derivable from ϕ if and only if ϕ ∨ ψ was derivable. Not just any connective has this intimate relation with derivability; it is a special feature of the conditional.

The Deduction Theorem is only a small taste of what lies ahead in metatheory. We will find as we go along that the mathematical perspective on logical reasoning that an abstract proof system opens up to us makes it possible to derive a wide array of very powerful, and often quite surprising, results. In the next section, we will develop a further metatheoretical result of considerably more complexity and interest than the Deduction Theorem.

A Completely Mechanized Proof Strategy [Next]

Earlier, I gave a number of strategical hints for constructing proofs in our proof system. While these hints will help guide you through many of the complexities of the proof process, they will not, in the more complex cases, tell you exactly what to do to successfully complete the proof—a measure of creativity and intellectual effort is still needed on the part of the human prover. In this section, I want to show that it is possible to specify a proof strategy that will do all the work for you—one which can be followed completely mechanically and which requires no creative or intellectual effort on the part of the human prover.

Three questions are immediately raised by this proposed project:

**Question #1:** If there's a completely mechanical proof strategy available, why did you waste my time giving all those imperfect strategical hints earlier? Why not give me the ideal strategy right away? It reminds me of how annoying it was in algebra class to go through all the work of learning to complete the square, only to then be taught the quadratic formula which would do all the work automatically.

**Answer:** Well, you're forced to learn to complete the square before learning the quadratic formula on the theory that completing the square gives the principles behind the quadratic formula and thus justifies you in using that formula. But there's a much more satisfying explanation of why you had to go through the "imperfect" proof strategies before getting to the current completely mechanized strategy—those earlier strategies, when they work (which is the great majority of the time) are vastly more efficient than the strategy we're going to give now. The current strategy will always work, but it will do so quite slowly. Proofs that can be done in a mere ten lines using the conventional strategies can easily stretch to a few hundred lines using this method. As with most things in life, the more certainty you want, the more you have to pay for it. The price might be worth it when you're being sure that you'll have health insurance in case of catastrophic illness, or being sure that your new airplane won't explode in midair, but it's really not worth it to avoid the slight risk of having to use your brain at some point in a proof.

**Question #2:** OK, if this new proof strategy is so time-intensive, why bother with it at all?

**Answer:** Because the point of this proof procedure isn't to allow us to construct proof using it. Rather, it's to establish certain general facts regarding the proof system. In particular, two features of the proof system will be brought out by thinking about this method:

1. **First,** it is possible to reliably implement the proof system on a computer. A strategy which at some point resorts to "if this doesn't work, see if you can find your own way to complete the proof" can't be made into a program, but the strategy that I'm about to give can quite easily be implemented in a program. This is interesting both on general theoretical grounds, insofar as it shows something about the ability of computers to simulate human cognitive processes, and on immediately practical grounds, because the computational efficiency issues mentioned in the answer to the previous question are less of an issue when the proof is being constructed by computer.

2. **Second,** even if we never use the mechanized proof strategy (and we never will use it, once we've set it out), the mere fact that it exists reveals certain general features of the proof system for sentential logic. The fact that there can be a mechanized proof strategy places certain upper bounds on the complexity of the total collection of logical theorems in sentential logic, for example. This is an issue we will come back to in much more detail later. And by considering the range of claims that can be proved using the mechanized proof strategy, we will be able to place a useful lower bound on the power of the proof system, bringing us a step closer to a full characterization of how much can be accomplished within that system. Again, this is an issue we will come back to later. In short, then, a mechanized proof strategy is a valuable step toward the realization of significant metatheoretic goals.

**Question #3:** What exactly does it mean to say that this proof strategy is completely mechanized?

**Answer:** This is a very good question. As a first stab, we might say that a proof strategy is completely mechanized if it gives a series of instructions which, if followed reliably, will always lead to the construction of the desired proof (should the claim in question actually be provable). However, there's an obvious difficulty with this take on what it is for a process to be completely mechanized. Suppose I give the following proof strategy:

**The Almost Completely Useless Proof Strategy**

\[ Φ \iff Θ \]

...
To prove $\Phi_1, \ldots, \Phi_n \vdash \Theta$

**Step 1:** For each $\Phi_i$, write a line in the proof of the form:

\[(i) \, \Phi_i \vdash A\]

**Step 2:** Add a Show line of the form:

\[(n + 1) \, \text{Show} \, \Theta\]

**Step 3:** Use the proof rules to produce a subsequent line, not imbedded in any subproof, which contains the sentence $\Theta$ as its content.

Now, in a sense this is a great proof strategy. Follow it, and you'll complete every valid proof successfully. However, in another, more significant, sense, it's almost completely useless. **Step 3** just has a bit too much built into it. It's true that if you do what **Step 3** tells you to do, you'll complete the proof, but the following of **Step 3** just doesn't seem like a mechanical process. It's too much to do.

One is tempted to say that the problem is that the strategy doesn't tell you how to carry out **Step 3**. But this can't be quite right. We can't expect a proof strategy to tell us how to carry out each of its steps, because this leads to an infinite regress. Suppose it were our (foolish) goal to have the strategy tell us how to perform each step. Suppose then that we had some $n$th step, which was of the form:

- **Step $n$:** Perform action $X$.

Now, in keeping with our intention to tell us how to do everything, we need further instructions telling us how to $X$. So we (let's say) add an $n$-plus-first step:

- **Step $n+1$:** Perform action $Y$ in order to achieve $X$.

That's great, but of course it leaves us wondering how to perform $Y$. So we add an $n$-plus-second step:

- **Step $n+2$:** Perform action $Z$ in order to achieve $Y$.

But then how do we $Z$? Well, **Step $n+3$** will tell us to $W$ in order to $Z$, and **Step $n+4$** will tell us to $V$ in order to $W$, and so on.

So the strategy can't tell us how to do everything, since that would require it to tell us an infinite amount. There have to be certain actions which are so basic that the strategy doesn't have to tell us how to do them. The problem with the Almost Completely Useless Proof Strategy, then, was really that it wasn't built up out of these basic actions -- **Step 3** was too complicated an action, and needs to be broken down into the primitive action vocabulary.

But now a new, and more interesting, question emerges: what counts as the primitive vocabulary? What sorts of actions are so simple and basic that an action recipe which leaves them unexplained counts as wholly mechanical? This is a difficult and substantial question, and it's not immediately clear how one would even begin to answer it. Suppose, for example, that I am setting out a wholly mechanical recipe for making lemon cheesecake ice cream. A step in the recipe like:

- Add enough sugar to balance out the bitterness of the lemons, while still leaving a pleasant hint of tartness.

is pretty plainly not wholly mechanical -- someone who read that recipe wouldn't automatically know exactly what to do. Suppose, then, that I modify it with:

- Add 1/2 cup sugar for every 2 tablespoons of lemon juice added earlier.

Is this rule wholly mechanical? Well, it's better than the previous rule, but it doesn't tell us how to know how much lemon juice we added earlier (is memory wholly mechanical?) or how to perform the computation to get the right amount of sugar (is multiplication wholly mechanical?). It also doesn't tell us how to find the sugar. Suppose, then, we add:

- Get sugar by locating the box labeled Sugar.

But is this good enough, or do we need rules for recognizing a box, and for recognizing the word Sugar (and perhaps then also for recognizing the letter S, and the letter u, and so on)?

Later we will return to this question in more detail, but for now I'm simply going to appeal to a rough and unfinished notion of wholly mechanical. We'll take as our rough-and-ready standard tasks that a computer can do, and in particular we will need the capacity to do each of the following:

- Recognize all of the symbols of our formal language.
- Recognize structural properties of formulas in the formal language, such as main connective and scope of a connective.
- Recognize correct input conditions for each of our proof rules, and determine the correct output given those inputs.
- Add new lines to a proof given a choice of proof rule and inputs to that proof rule.

Our goal, then, will be to show that there is a proof strategy each of whose steps involves only actions drawn from that list (or items sufficiently similar or otherwise primitive) which will always succeed in producing a proof.
With these questions out of the way, we're ready to turn to our main result. We will prove the following theorem:

**Theorem:** Given any sentences \( \Phi_1, \Phi_2, \ldots, \Phi_n \), and \( \Theta \) such that \( \Phi_1, \ldots, \Phi_n \models \Theta \), there is a proof of \( \Theta \) from the assumptions \( \Phi_1, \Phi_2, \ldots, \Phi_n \).

Notice first of all that this theorem says nothing about there being a mechanical proof procedure. It simply asserts that there is a proof of \( \Theta \) from the \( \Phi_n \), but says nothing about how to obtain that proof. I've omitted mention of the mechanical proof procedure because, as Question #3 above helps bring out, that notion is still too vague for us to prove theorems about. Given our only rough-and-ready notion of a mechanical procedure, I could no more prove that there is a mechanical procedure than I could prove that there is an *interesting* root of the equation \( x^2 - 7x - 44 \). Neither notion (mechanical or interesting) is sharply-defined for it to be possible to prove results about it. However, by looking at the way in which we prove the theorem, it will become obvious that I am setting out a proof procedure which meets our rough standard of being mechanical. We just can't say officially (i.e., in the content of the theorem itself) that we've got a mechanical procedure.

**Question:** But if the particular nature of the proof procedure is not part of the theorem, then surely the theorem itself is quite trivial, isn't it? All it says is that when some assumptions imply a conclusion, there's a proof of that conclusion from the assumptions. But that's just a definitional truth, isn't it? That's just what a proof is -- a way of showing that some assumptions imply a conclusion.

**Answer:** This is a very important point, so I want to take time to make sure it is perfectly clear. The theorem we are proving here is most definitely not a trivial one. The theorem links two quite different notions. One is the notion of implication -- a number of assumptions imply a conclusion just in case every interpretation that makes all of the assumptions true also makes the conclusion true. This is a semantic notion, one which centers around the semantic concept of truth in an interpretation. The other is the notion of derivability -- a conclusion can be derived from a number of assumptions just in case a proof can be constructed from the assumptions to the conclusion. This is a syntactic notion, which is built on a collection of syntactically specified proof rules. It is far from automatic that these two notions line up correctly. There are, after all, an infinite number of true implication claims, of arbitrarily great complexity. On the face of it, it's quite an audacious claim that the dozen-odd rules we've given for the proof system are enough to capture every single one of those implications. Of course, we would want a good proof system to do that capturing, but it remains to be shown (and this theorem will do a sizeable chunk of the showing) that our proof system *is* a good one.

So much for justifying the project of proving this theorem. Now on to the proof itself. Before launching into the details of the proof, let me emphasize that this proof will not be like the formal proofs in the sentential logic that we've been looking at above. The claim we are proving is one in English, about the logical system, rather than one in the logical system. (We could translate it into the logical system, but that would give us just \( P \), which clearly loses some of the interest of the claim.) Thus the proof I will give will also be in English, rather than a formal proof following our proof rules. It will thus be a more-or-less informal proof, of the style which is typically given for mathematical theorems. Later, we will see how to make completely formal proofs of this sort as well, but that's a long way down the road.

**Proof:** The proof of this theorem will proceed by establishing three lemmas, or small claims used in the proof of a larger theorem (you can think of lemmas as being like the results of subproofs in our formal proof system). The three lemmas in question are:

**Lemma 1:** Given sentences \( \Phi_1, \ldots, \Phi_n \), \( \Theta \) such that \( \Phi_1, \ldots, \Phi_n \models \Theta \), the disjunctive normal form of \( \Phi_1 \& \Phi_2 \& \ldots \& \Phi_n \& \neg \Theta \) is an inconsistent sentence.

**Lemma 2:** If \( \Phi \) is an inconsistent sentence in disjunctive normal form, then for every disjunct in \( \Phi \), there is some sentence letter \( \Psi \) such that both \( \Psi \) and \( \neg \Psi \) are conjuncts in that disjunct.

**Lemma 3:** Given any sentence \( \Phi \) and its disjunctive normal form \( \Phi^* \), there is a proof of \( \Phi^* \) from \( \Phi \).

Just to make sure the second lemma is clear, let's consider an example. Recall that a sentence is in disjunctive normal form if:

(a) It is a disjunction -- i.e., at the top level of logical form, it is of the form \( \Phi_1 \lor \Phi_2 \lor \ldots \lor \Phi_n \) for some \( n \) disjuncts.

(b) Each disjunct (that is, each \( \Phi_i \)) is itself a conjunction, and is thus of the form \( \Theta_1 \& \Theta_2 \& \ldots \& \Theta_m \).

(c) Each conjunct in that conjunction is either a sentence letter or a negation of a sentence letter.

Now, consider the following sentence in disjunctive normal form:

- \( (P \& \neg R) \lor (Q \& S \& \neg S \& \neg Q \& T) \lor (P \& \neg P) \lor (P \& Q \& R \& S \& T \& \neg S) \)

This is an inconsistent sentence -- there is no interpretation on which it comes out true (you can build a truth table for it if you doubt that). It has four disjuncts:

- \( P \& R \& Q \& \neg R \)
Each of these four disjuncts contains some explicit contradiction in the form of a sentence letter and the negation of that sentence letter. In the first disjunct, we find both \( R \) and \( \neg R \). In the second, we find both \( Q \) and \( \neg Q \) (or both \( S \) and \( \neg S \)). In the third, we find both \( P \) and \( \neg P \), and in the fourth both \( S \) and \( \neg S \). Lemma 2 claims that any inconsistent sentence in disjunctive normal form will be like this – each of its disjuncts will conjunct some sentence letter with its own negation. We don't have a proof for that yet – so far I've only asserted Lemma 2, not established it – but I first want to make sure it's clear what's being claimed.

Before trying to prove any of the three lemmas, I want to briefly assume they are true, and show how the main theorem will follow from these lemmas. I find it's easier to follow the small details of a proof if you know how those details are leading you to the final goal. Suppose for the moment, then, that all three lemmas are true. How would that help us? Well, suppose we've got some sentences \( \Phi_1, \Phi_2, \ldots, \Phi_n \) and \( \Theta \), such that \( \Phi_1, \ldots, \Phi_n \models \Theta \) (thus meeting the conditions of application for the theorem). We want to show that \( \Theta \) can be derived from the \( \Phi \)s. We can start our proof like this:

| (1) \( \Phi_1 \) | A |
| (2) \( \Phi_2 \) | A |
| ... |
| (n) \( \Phi_n \) | A |
| (n+1) Show \( \Theta \) |
| (n + 2) \( \neg \Theta \) | AIP |

We are thus beginning an indirect proof of \( \Theta \). To succeed, we need to derive some contradiction. To make things specific, we'll try to derive both \( P \) and \( \neg P \). The first thing we will do is collect all the \( \Phi \)s and the negation of \( \Theta \) into a big conjunction:

| (1) \( \Phi_1 \) | A |
| (2) \( \Phi_2 \) | A |
| ... |
| (n) \( \Phi_n \) | A |
| (n+1) Show \( \Theta \) |
| (n + 2) \( \neg \Theta \) | AIP |
| (n + 3) \( \Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta \) &I, 1,2,3,...,n,n+2 |

**Note:** I am stretching our original definition of \&I here, since \&I technically allows only two inputs to be joined into a conjunction. (I'm also stretching the syntax, since the precise syntax would require that the sentence on line \( n+3 \) have lots of parentheses – \( 2n + 2 \) of them – to indicate internal grouping, but we decided a long time ago that we would allow internal parentheses to be grouped in extended conjunctions and disjunctions.) In what follows, I will stretch our inference rules in the following ways:

- \&I will allow any (finite) number of inputs to be combined into an extended conjunction
- \&E will allow any conjunct of an extended conjunction to be inferred.
- \lor will allow any (finite) number of new disjuncts to be combined with an input to make an extended disjunction
- \lor E will allow an extended disjunction \( \psi_1 \lor \psi_2 \lor \ldots \lor \psi_n \) to be combined with a collection of conditionals \( \psi_1 \rightarrow \Omega, \psi_2 \rightarrow \Omega, \ldots, \psi_n \rightarrow \Omega \) to infer \( \Omega \).

Each of these extensions of the earlier rules can be justified, although each of the justifications requires an appeal to mathematical induction (because the extensions allow an indefinitely large number of inputs/output components), and thus fall outside the scope of the justifications we've done so far.

On line \( n+3 \) of our proof we have the extended conjunction \( \Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta \). By lemma 2, we know that we can derive from this extended conjunction its disjunctive normal form. We don't yet know how that derivation will go, but we can just stick in an ellipsis in the proof, to get:

| (1) \( \Phi_1 \) | A |
| (2) \( \Phi_2 \) | A |
| ... |
| (n) \( \Phi_n \) | A |
| (n+1) Show \( \Theta \) |
| (n + 2) \( \neg \Theta \) | AIP |
| (n + 3) \( \Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta \) &I, 1,2,3,...,n,n+2 |
| ... |
(n + m) \((\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \lor (\Phi_{2,1} \& \ldots \& \Phi_{2,n_2}) \lor \ldots \lor (\Phi_{k,1} \& \ldots \& \Phi_{k,n_k})\)

where \((\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \lor (\Phi_{2,1} \& \ldots \& \Phi_{2,n_2}) \lor \ldots \lor (\Phi_{k,1} \& \ldots \& \Phi_{k,n_k})\) is the disjunctive normal form of \(\Phi_1 \& \Phi_2 \& \ldots \& \Phi_n \& \neg \Theta\).

By assumption, \(\Phi_1, \ldots, \Phi_n \models \Theta\), so by lemma 1, the disjunctive normal form of \(\Phi_1 \& \Phi_2 \& \ldots \& \Phi_n \& \neg \Theta\) is inconsistent. By lemma 2, we then know that each disjunct of \((\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \lor (\Phi_{2,1} \& \ldots \& \Phi_{2,n_2}) \lor \ldots \lor (\Phi_{k,1} \& \ldots \& \Phi_{k,n_k})\) contains as two of its conjuncts some sentence letter and the negation of that sentence letter. The next step in the proof will then be to perform an extended dilemma reasoning, one in which we will prove one conditional for each branch of the extended disjunction. The fact that each disjunction contains a contradiction will be crucial to the success of this dilemma reasoning.

Let’s look at how one of these conditional proofs will proceed. Since the first disjunct in the disjunctive normal form is:

- \(\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}\)

we need to prove:

- \((\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \rightarrow (P \& \neg P)\)

(Recall that we decided earlier that we would use \(P\) and \(\neg P\) as the two halves of the contradiction we needed to finish the indirect proof.) Assume that \(\Phi_{1,j}\) and \(\Phi_{1,k}\) are the two conjuncts which are some sentence letter and its negation. Then we perform the following proof moves:

\[(p) \text{ Show } (\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \rightarrow (P \& \neg P)\]

\[(p + 1) \Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1} \text{ ACP}\]

\[(p + 2) \Phi_{1,j} \text{ &E, } p+1\]

\[(p + 3) \Phi_{1,k} \text{ &E, } p+1\]

\[(p + 4) P \& \neg P \text{ !, } p+2, p+3\]

Since \(\Phi_{1,j}\) and \(\Phi_{1,k}\) are some sentence letter and its negation, they form a contradiction and hence provide the needed input to the ! rule.

Note: If there is more than one contradictory pair in \(\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}\), we will extract the first one we come to use with the ! rule. This policy matters only for making sure the proof procedure is wholly mechanical.

We will follow a similar procedure for each of the disjuncts in the disjunctive normal form. Since each disjunct, by lemma 2, must contain some contradictory pair, there will always be appropriate \(\Phi_s\) to extract to trigger the use of the ! rule. Doing this for each disjunct, we get:

\[(1) \Phi_1 \text{ A}\]

\[(2) \Phi_2 \text{ A}\]

\[\ldots\]

\[(n) \Phi_n \text{ A}\]

\[(n+1) \text{ Show } \Theta\]

\[(n + 2) \neg \Theta \text{ AIP}\]

\[(n + 3) \Phi_1 \& \Phi_2 \& \ldots \& \Phi_n \& \neg \Theta \text{ &I, } 1,2,3,\ldots,n,n+2\]

\[\ldots\]

\n\n\n\n\[\text{[Insert here derivation of disjunctive normal form]}\]

\[(n + m) \text{ Show } (\Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1}) \rightarrow (P \& \neg P)\]

\[(n + m + 1) \Phi_{1,1} \& \Phi_{1,2} \& \ldots \& \Phi_{1,n_1} \text{ ACP}\]

\[(n + m + 3) \Phi_{1,j} \text{ &E, } n+m+2\]

\[(n + m + 4) \Phi_{1,k} \text{ &E, } n+m+2\]

\[(n + m + 5) P \& \neg P \text{ I, } n+m+3, n+m+4\]

\[(n + m + 6) \text{ Show } (\Phi_{2,1} \& \Phi_{2,2} \& \ldots \& \Phi_{2,n_2}) \rightarrow (P \& \neg P)\]

\[(n + m + 7) \Phi_{2,1} \& \Phi_{2,2} \& \ldots \& \Phi_{2,n_2} \text{ ACP}\]

\[(n + m + 8) \Phi_{2,j} \text{ &E, } n+m+7\]

\[(n + m + 9) \Phi_{2,k} \text{ &E, } n+m+7\]

\[(n + m + 10) P \& \neg P \text{ I, } n+m+8, n+m+9]\]

\[\ldots\]

\[\text{[Insert here similar proofs of conditionals for other disjuncts of the extended disjunction on line } n + m\]
Once we've done all of these conditional proofs, we're ready to use \( \lor E \) to derive \( P \land \neg P \) directly from the disjunctive normal form. This moves \( P \land \neg P \) up to the top level of the proof, and then we just extract \( P \) and \( \neg P \) to derive the two halves of a contradiction -- thereby completing the indirect proof and cancelling the original Show line of \( \Theta \). We then have:

\[
(1) \Phi_1 \quad \text{A}
(2) \Phi_2 \quad \text{A}
\ldots
(n) \Phi_n \quad \text{A}
\]

\[
(n+1) \text{Show } \Theta
(n+2) \neg \Theta
(n+3) \Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta \quad \text{&I}, 1,2,3,\ldots,n,n+2
\]

[Insert here derivation of disjunctive normal form]

\[
(n+m) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, AIP
\]

\[
(n+m+1) \text{Show } \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta \quad (P \land \neg P)
\]

\[
(n+m+2) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, ACP
\]

\[
(n+m+3) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+4) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+5) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+6) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+7) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+8) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+9) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+10) \Phi_1,1 \land \Phi_1,2 \land \ldots \land \Phi_{1,n1} \land \neg \Theta, \text{&I}, n+m+2
\]

\[
(n+m+5k-4) \text{Show } \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta
\]

\[
(n+m+5k-3) \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta, ACP
\]

\[
(n+m+5k-2) \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta, \text{&I}, n+m+5k-3
\]

\[
(n+m+5k-1) \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta, \text{&I}, n+m+5k-3
\]

\[
(n+m+5k) \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta, \text{&I}, n+m+5k-2
\]

\[
(n+m+5k+1) \Phi_k,1 \land \Phi_k,2 \land \ldots \land \Phi_{k,nk} \land \neg \Theta, \text{&I}, n+m+5k+1
\]

This is now a complete proof of \( \Theta \) from \( \Phi_1, \ldots, \Phi_n \).

The above considerations show that, given the correctness of lemmas 1 through 3, we can construct a proof of \( \Theta \) from the \( \Phi \)s. (Moreover, we can construct it in a wholly mechanical way, supposing that lemma 3 gives us a mechanical method of deriving the disjunctive normal form of \( \Theta \) from \( \Phi \)). Once we show that each of the three lemmas is correct, then, we will have established the theorem. It is to this task that we now turn.

**Lemma 1**: Given sentences \( \Phi_1, \ldots, \Phi_n, \Theta \) such that \( \Phi_1, \ldots, \Phi_n \models \Theta \), the disjunctive normal form of \( \Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta \) is an inconsistent sentence.

**Proof**: Suppose that \( \Phi_1, \ldots, \Phi_n \models \Theta \). We know that when a collection of premises entail some conclusion, then the set of those premises and the negation of the conclusion is inconsistent. This is because the entailment relation shows us that any interpretation making the premises true makes the conclusion true as well, and hence that there is no interpretation making the premises true and the conclusion false. But an interpretation making the conclusion false makes the negation of the conclusion true, so there is no interpretation making the premises and the negation of the conclusion all true. So, in particular, we know that:

* \( (\Phi_1 \land \Phi_2 \land \ldots \land \Phi_n \land \neg \Theta) \)

is an inconsistent set.
Now consider the sentence:

- \( \phi_1 \land \phi_2 \land \ldots \land \phi_n \land \neg \theta \)

Suppose there were some interpretation \( I \) making this sentence true. Then, by the nature of \( \land \), \( I \) would have to make true each of \( \phi_1, \phi_2, \ldots, \phi_n, \) and \( \neg \theta \). But since \( (\phi_1, \ldots, \phi_n, \neg \theta) \) is inconsistent, we know there is no such interpretation. Hence \( \phi_1, \ldots, \phi_n, \neg \theta \) is an inconsistent sentence.

We saw earlier that for every sentence in the language, there is some sentence in disjunctive normal form equivalent to that sentence. So there is some sentence -- call it \( (\phi_1 \land \ldots \land \phi_n \land \neg \theta)^* \) -- which is the disjunctive normal form of \( \phi_1 \land \ldots \land \phi_n \land \neg \theta \). Since \( (\phi_1 \land \ldots \land \phi_n \land \neg \theta)^* \) is equivalent to \( \phi_1 \land \ldots \land \phi_n \land \neg \theta \), and since \( \phi_1, \ldots, \phi_n, \neg \theta \) is inconsistent, we know that \( (\phi_1 \land \ldots \land \phi_n \land \neg \theta)^* \) is also inconsistent. This suffices to prove lemma 1.

**Lemma 2:** If \( \phi \) is an inconsistent sentence in disjunctive normal form, then for every disjunct in \( \phi \), there is some sentence letter \( \psi \) such that both \( \psi \) and \( \neg \psi \) are conjuncts in that disjunct.

**Proof:** Suppose that \( \phi \) is an inconsistent sentence in disjunctive normal form. Then every interpretation makes \( \phi \) false. Since \( \phi \) is an extended disjunction, it is true if any of its disjuncts are true. Thus each of \( \phi \)'s disjuncts must also be made false by every interpretation.

Consider some one of \( \phi \)'s disjuncts. It must be a conjunction of the form:

- \( \psi_1 \land \psi_2 \land \ldots \land \psi_n \)

for some \( n \), where each \( \psi_i \) is either a sentence letter or a negation of a sentence letter. We know that this disjunct must be false relative to every interpretation. Suppose that no sentence letter showed up in \( \psi_1 \land \ldots \land \psi_n \), both unnegated and negated. Then we can divide the \( \psi_i \)'s into two groups -- the unnegated and the negated -- with no sentence letter showing up in both groups. Let \( \psi^* \) be the set of unnegated \( \psi_i \)'s, and \( \psi^- \) be the set of negated \( \psi_i \)'s. Now pick an interpretation \( I \) such that \( I \) assigns \( T \) to all the sentence letters in \( \psi^* \) and \( F \) to all the sentence letters in \( \psi^- \). Since there is no overlap between \( \psi^* \) and \( \psi^- \), there must such an \( I \).

The importance of \( I \) is that it makes true the disjunct \( \psi_1 \land \ldots \land \psi_n \). It makes true every conjunct in \( \psi^* \), since each conjunct in \( \psi^* \) is an unnegated sentence letter, and \( I \) assigns \( T \) to all of those sentence letters. It makes true every conjunct in \( \psi^- \), since every conjunct in \( \psi^- \) is a negated sentence letter, and \( I \) assigns \( F \) to all of those sentence letters, yielding \( T \)'s after negation is applied. Every conjunct in \( \psi_1 \land \ldots \land \psi_n \) is in either \( \psi^* \) or \( \psi^- \), so every conjunct in \( \psi_1 \land \ldots \land \psi_n \) is assigned \( T \) by \( I \). Thus assigns \( T \) to the whole disjunct \( \psi_1 \land \ldots \land \psi_n \), and thereby also assigns \( T \) to \( \phi \) (by making true one of its disjuncts). But this contradicts the assumption that \( \phi \) is inconsistent. Thus there can be no such \( I \) -- but such an \( I \) is inevitable if \( \psi_1 \land \ldots \land \psi_n \) does not contain some sentence letter both unnegated and negated. Thus \( \psi_1 \land \ldots \land \psi_n \) must contain some such contradictory pair.

The considerations showing that \( \psi_1 \land \ldots \land \psi_n \) must contain a contradictory pair are wholly general, and show that each disjunct in \( \phi \) must contain its own contradictory pair. This suffices to prove lemma 2.

**Lemma 3:** Given any sentence \( \phi \) and its disjunctive normal form \( \phi^* \), there is a proof of \( \phi^* \) from \( \phi \).

**Proof:** To prove lemma 3, we will give a general recipe for constructing a proof from \( \phi \) to \( \phi^* \). This recipe will involve four phases:

- Eliminate all biconditionals from \( \phi \)
- Eliminate all conditionals from \( \phi \)
- Move all negations in \( \phi \) inward in scope until they have smallest possible scope, and eliminate any double negations
- Use the \&v rule to give all conjunctions small scope and all disjunctions large scope

We'll look at the details of each of these phases, and track a particular case through all four. For our example, we will use the sentence:

- \( (P \rightarrow Q) \land ((Q \lor R) \leftrightarrow \neg S) \land \neg (P \rightarrow \neg S) \)

This sentence comes from conjoining the assumptions and the negation of the conclusion in the following argument:

- \( P \rightarrow Q \land (Q \lor R) \leftrightarrow \neg S \land P \rightarrow \neg S \)

The particular example, of course, does nothing in itself to establish the general lemma, since that lemma asserts that we can find an appropriate derivation for any sentence. The particular example will be purely for purposes of illustration. It's not essential that we use a sentence that is a conjunction of the assumptions and the negation of the conclusion of some argument -- the procedure we give here will derive the disjunctive normal form of any sentence whatever. But since it's this sort of sentence we are interested in for the specific purposes of this theorem, we'll use an example of that form.
The first phase of deriving a disjunctive normal form is to eliminate all the biconditionals from our target sentence. A sentence in disjunctive normal form contains only the connectives $\&$, $\lor$, and $\neg$, so all conditionals and biconditionals have to be eliminated to reach the appropriate form. In order to perform this elimination, we will introduce a new derived proof rule $\leftrightarrow^*$, which transforms a biconditional into a conjunction of conditionals in both directions. The new rule has the following form:

$$\Phi \leftrightarrow \Theta \; 
\leftrightarrow^*:\; \begin{array}{c}
\Phi \leftrightarrow \Theta \\
\equiv\Theta \\
(\Phi \rightarrow \Theta) \land (\Theta \rightarrow \Phi)
\end{array}$$

This rule is easily justified in both directions. In the top-to-bottom direction, it requires only two applications of $\leftrightarrow$. The second phase brings us to the following point:

$$(\Phi \rightarrow \Theta) \land (\Theta \rightarrow \Phi)$$

Note that the new rule of $\leftrightarrow^*$ is a bidirectional rule, and it can thus be used via Replacement on any biconditional in a sentence, no matter how deeply imbedded. This feature of $\leftrightarrow^*$ will be crucial to our strategy.

The first step in deriving the disjunctive normal form, then, will be to use $\leftrightarrow^*$ on each biconditional in our original sentence $\Phi$. To make the procedure wholly mechanistic, we’ll adopt the arbitrary convention that we proceed from left to right when there are multiple biconditionals in $\Phi$. Since each application of $\leftrightarrow^*$ reduces the total number of biconditionals in $\Phi$ by 1, and since there must be a finite number of biconditionals to begin with, we are guaranteed that a finite number of applications of $\leftrightarrow^*$ will result in the elimination of all biconditionals from $\Phi$.

With our particular example, we thus have the following beginning to the derivation of the disjunctive normal form:

$$(P \lor Q) \land ((Q \lor R) \leftrightarrow \neg S) \land \neg(P \lor \neg S)$$

At the end of the first phase of the derivation, we have a sentence which has no biconditionals in it, and which thus has only $\&$, $\lor$, $\neg$, and $\rightarrow$. The second phase requires eliminating all of the conditionals (including any new conditionals created through the application of $\leftrightarrow^*$ in the first phase). These eliminations will be carried out using the $\rightarrow\lor$ rule. The $\rightarrow\lor$ rule, like the $\leftrightarrow^*$ rule, is bidirectional, so again we can, via Replacement, use it to eliminate conditionals buried deep within a sentence. So we follow the same procedure as in the first phase -- working left to right, we replace each conditional with a disjunction of the negation of the antecedent with the consequent.

Continuing with our particular example, the second phase brings us to the following point:

$$(P \rightarrow Q) \land ((Q \lor R) \leftrightarrow \neg S) \land \neg(P \rightarrow \neg S)$$

At the end of the second phase, then, we have derived a sentence which contains only the connectives $\&$, $\lor$, and $\neg$. All that remains now is to get these connectives into their proper positions. Achieving the proper position requires three
things. First, negations must appear only with scope over sentence letters. Second, conjunctions must appear only with scope over sentence letters and negations. Third, disjunctions must appear with largest possible scope.

In the third phase we will implement the first of these requirements, by moving all negations inward in scope until they have scope only over sentence letters. Given that biconditionals and conditionals have now been eliminated, there are only three "bad" cases to worry about:

(i) A negation has scope over a disjunction
(ii) A negation has scope over a conjunction
(iii) A negation has scope over another negation

To deal with case (i), we will use the rule \( \rightarrow \) \( \rightarrow \) to move the negation inward in scope past the disjunction. To deal with case (ii), we will use the rule \( \rightarrow \) \( \rightarrow \) to move the negation inward in scope past the conjunction. To deal with case (iii), we will use the rule \( \rightarrow \) \( \rightarrow \) to eliminate pairs of negations until no negation has scope over another. Since each of these procedures will move a negation at least one step closer in scope to the sentence letters, and since negations must initially be some finite number of steps away in scope from the sentence letters, a finite number of applications of \( \rightarrow \) \( \rightarrow \), \( \rightarrow \), \( \rightarrow \), \( \rightarrow \), \( \rightarrow \), and \( \rightarrow \) \( \rightarrow \) must result in all negations having scope only over sentence letters.

To maintain a wholly mechanical procedure, we will legislate that we will tackle negations moving from left to right, moving on to the next negation only when the current target negation has been moved inward as far as possible.

Let's look now at how this process plays out with our specific example. Continuing from where we left off, we obtain:

\[
\begin{align*}
(1) & \ (P \rightarrow Q) \land ((Q \lor R) \leftrightarrow \neg S) \land \neg (P \rightarrow \neg S) \\
(2) & \ (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land \neg (S \rightarrow (Q \lor R)) \land \neg (P \rightarrow \neg S) \leftrightarrow \rightarrow, 1 \\
(3) & \ (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (\neg P \rightarrow \neg S) \rightarrow \lor, 2 \\
(4) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (\neg P \rightarrow \neg S) \rightarrow \lor, 3 \\
(5) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 4 \\
(6) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 5 \\
(7) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 6 \\
(8) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 7 \\
(9) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 8 \\
(10) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (P \lor \neg S) \rightarrow \lor, 9 \\
(11) & \ (P \lor Q) \land (\neg (Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (P \lor \neg S) \rightarrow \lor, 10 \\
\end{align*}
\]

Notice that the first negation (that forming \( \neg P \)) doesn't require any treatment, since it already has scope only over a sentence letter. Our first move, then, is to push inward the negation in \( \neg (Q \lor R) \), to form \( \neg Q \land \neg R \). The negations in \( \neg Q \) and \( \neg R \) are acceptable, as is the negation, further to the right, in \( \neg S \). The double negation in \( \neg \neg S \) is eliminated using \( \neg \neg \), and then the negation in \( \neg \neg (P \lor \neg Q) \) is pushed inward using \( \neg \neg \) to form \( \neg \neg P \land \neg \neg Q \). Two applications of \( \neg \neg \) then eliminate these two double negations to bring us to step 11, in which all negations have scope only over single sentence letters.

This brings us to the fourth and final phase, in which we adjust the relative scopes of \( \land \) and \( \lor \). To reach a disjunctive normal form, we need to make sure that no conjunction has scope over a disjunction. To do this, we will apply the rule \( \land \lor \), which has the effect of taking a conjunction with scope over a disjunction and transforming it into a disjunction with scope over conjunctions. Each time we apply \( \land \lor \), some conjunction will be moved inside the scope of some disjunction. Since a conjunction can have scope over only finitely many disjunctions, it is guaranteed that finitely many applications of \( \land \lor \) will move any given conjunction to its appropriate position inside the scope of all disjunctions. Moving left to right, dealing with each conjunction in this manner, we will eventually move all conjunctions to their appropriate position and obtain a sentence in disjunctive normal form.

Here's how it works out with our specific example. Continuing from where we left off, we obtain:

\[
\begin{align*}
(1) & \ (P \rightarrow Q) \land ((Q \lor R) \leftrightarrow \neg S) \land \neg (P \rightarrow \neg S) \\
(2) & \ (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land \neg (S \rightarrow (Q \lor R)) \land \neg (P \rightarrow \neg S) \leftrightarrow \rightarrow, 1 \\
(3) & \ (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (\neg P \rightarrow \neg S) \rightarrow \lor, 2 \\
(4) & \ (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (\neg P \rightarrow \neg S) \rightarrow \lor, 3 \\
(5) & \ (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 4 \\
(6) & \ (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 5 \\
(7) & \ (P \rightarrow Q) \land ((Q \lor R) \lor \neg S) \land (\neg S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 6 \\
(8) & \ (P \rightarrow Q) \land ((Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 7 \\
(9) & \ (P \rightarrow Q) \land ((Q \lor R) \lor \neg S) \land (S \lor Q \lor R) \land (\neg P \lor \neg S) \rightarrow \lor, 8 \\
\end{align*}
\]
The sentence on line 21 is, at long last, in disjunctive normal form. (Don’t worry too much, by the way, if you didn’t follow all the details of applying &\& in this last phase — I’m not so sure I got it all right myself. The important point is that you see how &\& can be used to move all of the conjunctions inside the disjunctions, even if you aren’t sure you could make it work out in a complicated example.)

The final disjunctive normal form we derive for (P \rightarrow Q) & ((Q \lor R) \leftrightarrow \neg S) & \neg(P \rightarrow \neg S) has twelve disjuncts. As predicted by lemma 2, each of these disjuncts contains some sentence letter and its negation. Here are the twelve disjuncts, with a contradiction indicated in red in each one.

- P & \neg Q & \neg R & S & P & Q
- P & \neg Q & \neg R & Q & P & S
- P & \neg Q & \neg R & R & P & S
- P & \neg S & S & P & S
- P & \neg S & Q & P & S
- P & \neg S & R & P & S
- P & S & Q & \neg Q & \neg R & R
- P & S & Q & \neg R & R & Q
- P & S & Q & \neg Q & \neg R & R
- P & S & Q & \neg R & R & Q
- P & S & Q & \neg Q & \neg R & R
- P & S & Q & \neg Q & \neg R & R

The table below illustrates the process of simplifying the disjunctive normal form.
We've now proved all three lemmas. Since we've already seen how to combine the three lemmas to produce a proof of the main theorem, this also suffices to prove the theorem. To see how it works out when all the pieces are put together, here's the proof that our proof procedure will generate for the following implication claim:

- \( P \rightarrow (Q \lor R) \leftrightarrow \neg S \) → \( P \rightarrow \neg S \)

(It's a long proof, if you want you can skip it.)

\[
\begin{array}{c|c}
1 & P \rightarrow Q \\
2 & (Q \lor R) \leftrightarrow \neg S \\
3 & Show P \rightarrow \neg S \\
4 & (P \rightarrow \neg S) \\
5 & (P \rightarrow Q) \land (Q \lor R) \leftrightarrow \neg S \land (P \rightarrow \neg S) \\
6 & (P \rightarrow Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
7 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
8 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
9 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
10 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
11 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
12 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
13 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
14 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
15 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
16 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
17 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
18 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
19 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
20 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
21 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
22 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
23 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
24 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S) \\
25 & (P \lor Q) \land ((Q \lor R) \rightarrow \neg S) \land (\neg S \rightarrow (Q \lor R)) \land (P \rightarrow \neg S)
\end{array}
\]
\[ \neg S \wedge Q \wedge \neg S \wedge R \]

(26) Show \( (\neg P \wedge \neg Q \wedge \neg R \wedge S \wedge P \wedge Q) \rightarrow (P \wedge \neg P) \)

(27) \( \neg P \wedge \neg Q \wedge \neg R \wedge S \wedge P \wedge Q \)

ACP

(28) \( \neg P \)

&E, 27

(29) \( P \)

&E, 27

(30) \( P \wedge \neg P \)

1, 28, 29

(31) Show \( (\neg P \wedge \neg Q \wedge \neg R \wedge Q \wedge P \wedge S) \rightarrow (P \wedge \neg P) \)

(32) \( \neg P \wedge \neg Q \wedge \neg R \wedge Q \wedge P \wedge S \)

ACP

(33) \( \neg P \)

&E, 32

(34) \( P \)

&E, 32

(35) \( P \wedge \neg P \)

1, 33, 34

(36) Show \( (\neg P \wedge \neg Q \wedge \neg R \wedge R \wedge P \wedge S) \rightarrow (P \wedge \neg P) \)

(37) \( \neg P \wedge \neg Q \wedge \neg R \wedge R \wedge P \wedge S \)

ACP

(38) \( \neg P \)

&E, 37

(39) \( P \)

&E, 37

(40) \( P \wedge \neg P \)

1, 38, 39

(41) Show \( (\neg P \wedge \neg S \wedge S \wedge P \wedge S) \rightarrow (P \wedge \neg P) \)

(42) \( \neg P \wedge \neg S \wedge S \wedge P \wedge S \)

ACP

(43) \( \neg P \)

&E, 42

(44) \( P \)

&E, 42

(45) \( P \wedge \neg P \)

1, 43, 44

(46) Show \( (\neg P \wedge \neg S \wedge Q \wedge P \wedge S) \rightarrow (P \wedge \neg P) \)

(47) \( \neg P \wedge \neg S \wedge Q \wedge P \wedge S \)

ACP

(48) \( \neg P \)

&E, 47

(49) \( P \)

&E, 47

(50) \( P \wedge \neg P \)

1, 48, 49

(51) Show \( (\neg P \wedge \neg S \wedge R \wedge P \wedge S) \rightarrow (P \wedge \neg P) \)

(52) \( \neg P \wedge \neg S \wedge R \wedge P \wedge S \)

ACP

(53) \( \neg P \)

&E, 52

(54) \( P \)

&E, 52

(55) \( P \wedge \neg P \)

1, 53, 54

(56) Show \( (P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge S) \rightarrow (P \wedge \neg P) \)

(57) \( P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge S \)

ACP

(58) \( Q \)

&E, 57

(59) \( \neg Q \)

&E, 57

(60) \( P \wedge \neg P \)

1, 58, 59

(61) Show \( (P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge Q) \rightarrow (P \wedge \neg P) \)

(62) \( P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge Q \)

ACP

(63) \( Q \)

&E, 62

(64) \( \neg Q \)

&E, 57

(65) \( P \wedge \neg P \)

1, 63, 64

(66) Show \( (P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge R) \rightarrow (P \wedge \neg P) \)

(67) \( P \wedge S \wedge Q \wedge \neg Q \wedge \neg R \wedge R \)

ACP

(68) \( Q \)

&E, 67

(69) \( \neg Q \)

&E, 67

(70) \( P \wedge \neg P \)

1, 68, 69

(71) Show \( (P \wedge S \wedge Q \wedge \neg S \wedge S) \rightarrow (P \wedge \neg P) \)

(72) \( P \wedge S \wedge Q \wedge \neg S \wedge S \)

ACP

(73) \( S \)

&E, 72

(74) \( \neg S \)

&E, 72

(75) \( P \wedge \neg P \)

1, 73, 74

(76) Show \( (P \wedge S \wedge Q \wedge \neg S \wedge Q) \rightarrow (P \wedge \neg P) \)

(77) \( P \wedge S \wedge Q \wedge \neg S \wedge Q \)

ACP

(78) \( S \)

&E, 77

(79) \( \neg S \)

&E, 77

(80) \( P \wedge \neg P \)

1, 78, 79

(81) Show \( (P \wedge S \wedge Q \wedge \neg S \wedge R) \rightarrow (P \wedge \neg P) \)
The procedure gets it done, but it certainly isn't quick about it. Compare the above monstrosity with the following much easier way of completing the same proof:

<table>
<thead>
<tr>
<th>(82)</th>
<th>P &amp; S &amp; Q &amp; ¬S &amp; R</th>
<th>ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(83)</td>
<td>S</td>
<td>&amp;E, 82</td>
</tr>
<tr>
<td>(84)</td>
<td>¬S</td>
<td>&amp;E, 82</td>
</tr>
<tr>
<td>(85)</td>
<td>P &amp; ¬P</td>
<td>I, 83,84</td>
</tr>
<tr>
<td>(86)</td>
<td>P &amp; ¬P</td>
<td>∨E, 25,26,31,36,41,46,51,56,61,66,71,76,81</td>
</tr>
<tr>
<td>(87)</td>
<td>P</td>
<td>&amp;E, 86</td>
</tr>
<tr>
<td>(88)</td>
<td>¬P</td>
<td>&amp;E, 86</td>
</tr>
</tbody>
</table>

A savings of 81 lines, and that's without even taking into account the compression that the derived rules build into the mechanical proof. (If you're a masochistic kind of person, and you've got far too much time on your hands, you might try to figure out how long the mechanical proof would be if it were done entirely in the core rule set. I don't know myself, but I'm guessing well over a thousand lines.)

Now that the theorem has been proved, let's step back for a moment and think about what it shows us. First, it show us that any time we've got some premises \( \Phi_1, \ldots, \Phi_n \) which imply (i.e., semantically imply) a conclusion \( \Theta \), then there is a proof of \( \Theta \) from these premises. This result is a sort of validation of our proof system, showing that provability is not just some arbitrary syntactic relation, but matches up with our starting notion of semantic logical implication. The theorem is, then, a start toward showing that we've got a good proof system, not just any old way of moving from one sentence to another.

However, this theorem is only a start on showing this, for two reasons:

- First, while the theorem goes a considerable distance toward showing that we can prove the things we want to in the proof system, it does nothing to show that we can't prove things we don't want to be able to prove. It's perfectly compatible with everything that was said in the proof of the theorem that it's possible to produce a proof of absolutely any claim whatsoever. This looks pretty unlikely -- how, for example, would you even begin to produce a proof of \( Q \) from \( P \)? -- but we don't have any way to rule it out. There are, after all, an infinite number of proofs that start from the assumption \( P \), and it's not clear how we can tell in advance that none of them, no matter how many millions of lines long, ever succeeds in proving \( Q \).

- Second, while the theorem shows that any time a finite number of premises \( \Phi_1, \ldots, \Phi_n \) imply a conclusion \( \Theta \), there is a proof of \( \Theta \) from the \( \Phi \)s, it leaves open the possibility that there are cases of implication in which a conclusion \( \Theta \) is implied by an infinite number of premises and in which there is no proof of \( \Theta \) from the \( \Phi \)s. The proof cannot block this possibility because if there are an infinite number of \( \Phi \)s, we cannot gather them together into a single conjunction (since sentences must be finite in length). To prove a perfect match between semantic implication and syntactic proof theory, then, we would need to show that implication is essentially finite -- that there are no unavoidably infinite consequence relations. This is in fact true, and we will prove it later.

One final question before we move on. What happens if we apply our wholly mechanical proof strategy to a case in which the assumptions do not in fact imply the conclusion? For example, what if we apply the strategy to:

\[ P \rightarrow Q, Q, \vdash P \]

Will we produce a proof of this claim? Well, we can certainly start the indirect proof by assuming \( ¬P \) via \( \text{AIP} \), and then form the conjunction of \( ¬P \) with the assumptions to get:

\[ (P \rightarrow Q) \& Q \& ¬P \]

Lemma 3 tells us that we can derive the disjunctive normal form of any sentence, so we'll be able to derive the disjunctive normal form of this sentence. If each disjunct of the normal form contains a contradictory pair, then we will be able to complete the proof in the manner indicated above. If, on the other hand, there is some disjunct of the normal form that does not contain a contradictory pair, then the proof procedure will not yield a completed proof (since the proof procedure looks for the contradictory pair, and will simply come to a halt if one is not found).

These considerations show the following:
If \( \Phi_1, \ldots, \Phi_n : \Theta \) is an argument (maybe a valid one, and maybe not), then the proof procedure will yield a proof of this argument if and only if the disjunctive normal form of \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) contains a contradictory pair in each disjunct.

We'd like to show that the following claim is true:

- **Claim:** If \( \Phi_1, \ldots, \Phi_n : \Theta \) is an invalid argument, then the proof procedure fails to produce a proof.

Given the conclusion we just reached, we can now see that this is equivalent to the following claim:

- If \( \Phi_1, \ldots, \Phi_n : \Theta \) is an invalid argument, then at least one disjunct of the disjunctive normal form of \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) does not contain a contradictory pair.

So, is this claim true? Well, it turns out that nothing we've said so far settles the issue. Lemmas 1 and 2 put together yield the following claim:

- If \( \Phi_1, \ldots, \Phi_n : \Theta \) is a valid argument, then every disjunct of the disjunctive normal form of \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) contains a contradictory pair.

This claim is equivalent to its contrapositive (the claim \( \neg Q \rightarrow \neg P \) corresponding to \( P \rightarrow Q \), which is:

- if some disjunct of the disjunctive normal form \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) does not contain a contradictory pair, then \( \Phi_1, \ldots, \Phi_n : \Theta \) is an invalid argument.

This is almost the claim that we want, but not quite -- its conditional is going in the wrong direction. Abstractly, lemmas 1 and 2 together yield a claim of the form \( P \rightarrow Q \), which is, via contraposition, equivalent to \( \neg Q \rightarrow \neg P \), but we want \( \neg P \rightarrow \neg Q \). If we could just strengthen the original conditional in \( P \rightarrow Q \) to a biconditional, then we could get what we wanted.

To get a biconditional -- that is, to get:

- if \( \Phi_1, \ldots, \Phi_n : \Theta \) is a valid argument if and only if every disjunct of the disjunctive normal form of \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) contains a contradictory pair.

we would need both lemma 1 and lemma 2 to be in the form of a biconditional. As they were originally stated, each is only a conditional:

- **Lemma 1:** Given sentences \( \Phi_1, \ldots, \Phi_n, \Theta \) such that \( \Phi_1, \ldots, \Phi_n \models \Theta \), the disjunctive normal form of \( \Phi_1 \& \Phi_2 \& \ldots \& \Phi_n \& \neg \Theta \) is an inconsistent sentence.

- **Lemma 2:** If \( \Phi \) is an inconsistent sentence in disjunctive normal form, then for every disjunct in \( \Phi \), there is some sentence letter \( \Psi \) such that both \( \Psi \) and \( \neg \Psi \) are conjuncts in that disjunct.

To make them biconditionals -- and thus to get what we need to establish our Claim -- we need to prove the conditionals in the other directions:

- **Lemma 1**: If \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) is an inconsistent sentence, then \( \Phi_1, \ldots, \Phi_n, \Theta \) is a valid argument.

- **Lemma 2**: If \( \Phi \) is a sentence in disjunctive normal form such that for every disjunct in \( \Phi \), there is some sentence letter \( \Psi \) such that both \( \Psi \) and \( \neg \Psi \) are conjuncts in that disjunct, then \( \Phi \) is inconsistent.

Each of these modified lemmas is easily established. To prove lemma 1”, note that if \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) is inconsistent, then every interpretation makes it false. Assume that there is an interpretation making \( \Phi_1, \ldots, \Phi_n \) true. Then to make \( \Phi_1 \& \ldots \& \Phi_n \& \neg \Theta \) false, it must make \( \neg \Theta \) false, which requires making \( \Theta \) true. But this is the definition of implication, so \( \Phi_1, \ldots, \Phi_n, \Theta \) is a valid argument. To prove lemma 2”, assume that \( \Phi \) is a sentence in disjunctive normal form with a contradictory pair in each disjunct. Now take an arbitrary interpretation. Clearly I cannot make true both members of a given contradictory pair, so I must make false at least one member of every such pair. Since that member will be a single conjunct in an extended conjunction, and since a single falsehood is enough to make false a whole conjunction, it follows that I will make false each disjunct in \( \Phi \). But a disjunction in which all the disjuncts are false is itself false, so I makes \( \Phi \) false. Since I was arbitrary, it follows that every interpretation makes \( \Phi \) false, and hence that \( \Phi \) is inconsistent.

So we see that our proof procedure is in fact a two-way test. Given any argument, whether valid or not, we can use the proof procedure to answer the question "Is the argument a valid one?". We simply start running the proof procedure on it, we will eventually be given one of two answers:

- Yes, the argument is valid. This answer is returned if the proof procedure generates a proof, which will happen only if the disjunctive normal form derived has a contradictory pair in every disjunct.

- No, the argument is not valid. This answer is returned if the proof procedure does not generate a proof, which will happen only if the disjunctive normal form has some disjunct without a contradictory pair.

Moreover, the answer returned by the proof procedure will definitely be a correct one. It will say that an argument is valid only if it really is valid, and say that an argument is invalid only if it really is invalid.

This is the best possible outcome for our question-answering procedure, and one that we could easily have fallen short of achieving. Compare, for example, this testing procedure with a hypothetical medical test for the presence of some illness. The best possible outcome, obviously, would be that we send a blood sample off to the lab and the lab always sends back an answer, yes or no, and that answer is always accurate. In real life, however, things rarely turn out so well. First, the test may yield false results. False results can come in two forms:

- False positive: the test returns an answer of "yes" when the correct answer is "no".

- False negatives: the test returns an answer of "no" when the correct answer is "yes".
Real medical tests, for example, almost always have some rate of both false positives and false negatives. The best one can hope for is that the rate of false positives and negatives stay low. But even very low rates of false results make it impossible to know for sure what the real answer to the test is, and they can even make probabilistic conclusions hard to draw accurately. Suppose, for example, that we have a highly accurate test for a rare medical condition X, which is found in roughly 1 out of every 10,000 in the population at large. Our test returns a false positive only 1% of the time. Suppose now that we test random people until we collect 100 positive results. Out of this 100, how many should we expect to be false positives?

It's tempting to say 1, but this turns out to be quite wrong. One false positive would leave 99 true positives. Given the prevalence rate of the disease in the population, we should expect to have to test 990,000 people to find 99 true positives. But in testing these 990,000 people with a test with a 1% chance of a false positive, we should expect 9900 false positives to show up. So we should expect, in testing this many people, to get (on average) a total of 9999 positives, of which only about 1% are true positives. Projecting these numbers to our original case, we can see that of our 100 positive results, we should expect that 99 are false positives. This is a highly counterintuitive result (and potentially a dangerous one, if those 100 people are started on treatment for condition X), and shows how easily a test which is less than perfectly accurate can become almost useless.

Second, even if the test never returns a false result, it might fail to return at all. Perhaps in certain kinds of cases, the laboratory has to keep testing and testing to see what the result is. Suppose, for example, that the test for condition X involves treating a tissue sample from the patient to low concentrations of chemical Y. Should reaction Z occur, the test would return a "yes" answer. Should reaction Z not occur, however, the test moves to an even lower concentration of Y. Since there's no limit to how low a concentration can be made, this test will never be able to return a "no" answer. So now consider the situation of a doctor who sends a sample to have this test performed on it. Perhaps after a time a "yes" answer will be sent back from the lab, and then he will know that the patient has condition X (since we're now assuming the test gives no false results). But suppose after a set amount of time (a day, let's say) the lab has not yet returned a result. The doctor can't conclude anything – it could be that the patient has condition X, and the lab just hasn't yet reached a sufficiently low concentration of Y. Or it could be that the patient does not have condition X, and the lab will never be able to reach a sufficiently low concentration of Y. There's no way the doctor can tell which of these is the case, so he can't tell what the status of his patient is. And no matter how long he waits, his situation will be no better (unless the lab does return a result) – if he wait a hundred years and the lab is still working, he still doesn't know. It's consistent with that that the lab is about to get a positive result, and also consistent that the lab never will get a positive result.

Here, then, is a list of ways in which a test can go bad:
- Sometimes produces false positives
- Sometimes produces false negatives
- Sometimes a true positive will cause the testing process never to come to an end.
- Sometimes a true negative will cause the testing process never to come to an end.

A particularly ill-conceived test might go wrong in all four of these ways. Imagine, for example, a test for medical condition X which proceeds like this: the doctor places a blank piece of paper on a table and waits for either the word "yes" or the word "no" to appear on it. This test could produce a false positive or a false negative (if, for example, a bored lab technician scrawls an answer at random on the paper, or if a quantum fluctuation causes some ink molecules to arrange themselves appropriately on the paper), and also could cause the testing process never to come to an end (as the piece of paper remains stubbornly blank). Perhaps for this reason such testing processes are rarely used in the medical profession. The testing process for correctness of an implication claim via the attempt to construct a proof using the wholly mechanical proof procedure, on the other hand, suffers from none of these deficiencies. The proof construction process will always terminate in a finite number of steps, and when it does it will infallibly indicate whether the argument at hand is a valid one.

**Question:** Given that following the mechanical proof procedure always produces reliable "no" answers when the target argument is an invalid one, doesn't this show that the earlier worry that our proof procedure might prove claims we don't want proved (because they are not genuine implications) has been resolved, and thus that our proof procedure neither undergenerates nor overgenerates?

**Answer:** Unfortunately, no. Although an attempt to construct a proof for an invalid argument using the mechanical proof procedure will always fail, this does nothing to show that other ways of trying to produce proofs of invalid claims will not succeed. There are still an infinite number of potential proofs of a particular invalid claim -- showing that one of this infinite number doesn't work still leaves quite a few to go. While we have in the mechanical proof procedure a reliable test for validity and invalidity of arguments, it is a separate task (one to which we'll return later) to show that the proof system as a whole is always a reliable guide to validity.

**Alternative Proof Systems [Next]**

The proof system that I have given here for sentential logic is by no means the only possible such system. I've given the system that I have because I think that it has a number of clear advantages over other methods of setting up a proof.
system, but now I want to take some time to look at some of those alternative methods. This discussion is, admittedly, a bit off the beaten track, and those who don't have a specific interest in proof systems may want to skip ahead to more core issues. However, I do think there are some good reasons to give a survey of proof system options. First, by looking at other systems, we can come to appreciate more fully some of the strengths of the system I've given here. Second, we'll come across some interesting advantages that some of the other proof systems have (advantages which we can keep in our bag of tricks to exploit later, should it prove necessary). Third, we'll get a better sense of what goes into designing a proof system, by seeing it done multiple times in various different ways. Fourth, there's a definite utility in being familiar with a variety of proof systems, especially if you plan to do work in logic. Logicians are notoriously diverse in the systems they use, so knowing many ways to set up a proof system can be as useful for a logician as knowing many languages can be for a well-traveled European. Fifth and finally (and this one is, admittedly, perhaps a reflection of a personal peculiarity), learning new proof systems is just kind of fun in itself.

We will start our tour of proof systems close at home, looking at rather small variants on the system given here. We'll then move outward, looking at increasingly unusual ways of setting up a proof system. I'm not going to cover every method of doing proofs that can be found in the literature (life is too short), but I'll try to hit all the major alternatives.

Variations on the Core Rules
The proof system I've given here has a total of eight substantive proof rules (in addition to the procedure rules A, ACP, AIP, and R), and three proof strategies. On the basis of these eight rules and three strategies, we were able to define and justify a large number of derived proof rules, and indefinitely many more such rules could also be introduced. Let's focus for the moment on just one of these rules -- that of $\lor E^*$, which has the following form:

$$
\lor E^*:\quad \Phi \lor \Theta \quad \text{or} \quad \neg \Phi \lor \neg \Theta
$$

Suppose that we had not included the rule $\lor E$ in our original proof system, but instead had used the rule $\lor E^*$. What difference would this have made in the proof system?

Answering this question requires saying what sort of differences we are interested in. Of course, with $\lor E^*$ replacing $\lor E$, various proofs would look different. Consider, for example, the claim:

• $P \lor Q, \neg Q, \therefore P$

In the original system (not using any derived rules), this claim would receive the following proof:

\begin{tabular}{ll}
(1) & $P \lor Q$ A \\
(2) & $\neg P$ A \\
(3) & Show Q \\
(4) & Show $P \rightarrow Q$ \\
 & \\
 & (5) & $P$ ACP \\
 & (6) & Show Q \\
 & (7) & $\neg Q$ AIP \\
 & (8) & $P$ R, 5 \\
 & (9) & $\neg P$ R, 2 \\
(10) & Show $Q \rightarrow Q$ \\
(11) & $Q$ ACP \\
(12) & $Q$ $\lor E$, 1,4,10
\end{tabular}

If, on the other hand, our proof system had $\lor E^*$ as its core rule of disjunction elimination, the proof would have looked like this:

\begin{tabular}{ll}
(1) & $P \lor Q$ A \\
(2) & $\neg P$ A \\
(3) & Show Q \\
(4) & $Q$ $\lor E^*$, 1,2
\end{tabular}

Clearly, then, there would be something different about the two proof systems. However, this is not a particularly deep difference. Although the two proofs are different in their mechanics, they still reach the same conclusion -- each proof system is in fact able to produce a proof of $Q$ from $P \lor Q$ and $\neg P$. 
We'll say that two proof systems are equal in power if they prove exactly the same theorems. Given a proof system P, we will write:
• \( \Phi_1, \ldots, \Phi_n \vdash P_\Theta \)
to mean that P produces a proof of \( \Theta \) from \( \Phi_1, \ldots, \Phi_n \). We will call the proof system we've given above \( \mathcal{P} \). Let's call the system with \( \lor E \) replaced by \( \lor E^* \), \( \mathcal{P}^* \). What we want to know, then, is whether \( \mathcal{P} \) and \( \mathcal{P}^* \) are equal in power -- that is, whether:
• Given any \( \Phi_1, \ldots, \Phi_n, \Theta \),
  \[ \Phi_1, \ldots, \Phi_n \vdash \mathcal{P} \Theta \]
  if and only if \( \Phi_1, \ldots, \Phi_n \vdash \mathcal{P}^* \Theta \)

One direction of this biconditional we've already effectively established. Suppose we've got some proof in \( \mathcal{P}^* \) which derives \( \Theta \) from \( \Phi_1, \ldots, \Phi_n \). We've already seen that \( \lor E^* \) is a derivable rule within \( \mathcal{P} \), so any use of \( \lor E^* \) can be replaced by a series of steps using only the core rules. So whatever proof we have in \( \mathcal{P}^* \), we can create a new proof in \( \mathcal{P} \) of the same result just by "defining away" the uses of \( \lor E^* \).

The other direction of the biconditional, however, is still an open question. For all we know there are results that can be derived within \( \mathcal{P} \) that cannot be derived within \( \mathcal{P}^* \). We will now show that this is not the case -- that every result provable in \( \mathcal{P}^* \) is also provable in \( \mathcal{P}^* \). Since \( \mathcal{P}^* \) has all the proof rules of \( \mathcal{P} \) except for \( \lor E \), the only possible difficulty lies in proofs that use \( \lor E \). To show that \( \mathcal{P}^* \) is equal in power to \( \mathcal{P} \), then, we need to show that \( \mathcal{P}^* \) can reproduce the effects of an application of \( \lor E \) using its rules. We need, that is, to show that \( \lor E \) is a derivable rule within \( \mathcal{P}^* \).

Suppose, then, that we've got three assumptions \( \Phi \lor \Theta, \Phi \rightarrow \psi, \Theta \rightarrow \psi \). \( \lor E \) will allow us to derive \( \psi \) from these three assumptions. Can we also derive \( \psi \) in \( \mathcal{P}^* \) yes. The necessary proof proceeds as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>( \Phi \lor \Theta )</td>
</tr>
<tr>
<td>(2)</td>
<td>( \Phi \rightarrow \psi )</td>
</tr>
<tr>
<td>(3)</td>
<td>( \Theta \rightarrow \psi )</td>
</tr>
<tr>
<td>(4)</td>
<td>Show ( \psi )</td>
</tr>
<tr>
<td>(5)</td>
<td>( \neg \psi )</td>
</tr>
<tr>
<td>(6)</td>
<td>Show ( \neg \Phi )</td>
</tr>
<tr>
<td>(7)</td>
<td>( \Phi )</td>
</tr>
<tr>
<td>(8)</td>
<td>( \psi )</td>
</tr>
<tr>
<td>(9)</td>
<td>( \neg \psi )</td>
</tr>
<tr>
<td>(10)</td>
<td>( \Theta \lor E^*, 1,6 )</td>
</tr>
<tr>
<td>(11)</td>
<td>( \psi \lor E, 3,10 )</td>
</tr>
</tbody>
</table>

\( \mathcal{P}^* \), then, can do everything that \( \mathcal{P} \) can do, and vice versa. Assuming (as we'll show definitively later, and as we got a good start on showing in the last section) that \( \mathcal{P} \) does what we want a proof system to do, it follows that \( \mathcal{P}^* \) also does what we want a proof system to do, and hence that it is also a live option for such a system.

What we want to look at in this section, then, is what other collections of rules we could have used to produce a proof system equal in power to \( \mathcal{P} \). We'll look first at several major choice points one encounters in setting up a system along the lines of \( \mathcal{P} \), and then consider some rule systems which diverge more widely from \( \mathcal{P} \). There are an infinite number of ways that one can set up a system of rules equal in power to \( \mathcal{P} \) (this is a substantive claim, by the way, so don't just take my word for it -- think about it and see if you can see why this is true), so we won't be able to consider all of the outliers, but we'll try to get a representative sample (enough, at least, to give a feel for the considerations that go into assembling an adequate collection of rules).

Proofs With Dependency Tracking

Proofs With Sequents

Proofs in an Axiomatic System