A Guide to Good Reasoning

By

Travis Figg

A Note to the Reader

I taught philosophy as an adjunct instructor for several years, including many courses in Critical Thinking. I gradually became frustrated with the various critical thinking texts which I had examined and eventually decided to make my own. However, I was only able to use this book in a couple of courses before I left academia. Rather than simply let it languish on my computer, which it has done since 2020, I decided it would be better to make it available to others. I would be delighted if instructors use it in their courses, or by anyone interested in becoming a better critical thinker. This is technically a draft instead of a polished manuscript, but I do not really have any intentions of revising it.

Table of Contents

Part I. Introduction

Chapter 1. Introducing Critical Thinking

You, like all conscious human beings, have a great many beliefs. These include beliefs about who you are, what your interests are, who your family members are, who your friends are, what kinds of food you like, and on and on. It would be exhausting to list all the different things you believe.

What is more, you make decisions based on what you believe. If you want to eat at a certain restaurant for dinner, you travel in the direction which you believe the restaurant to be, and you leave so you have time to get there before you believe it closes. If you want to make a meal yourself, you look until you find what you believe to be a good recipe. If you want to cheer up a friend after a bad day, you do something that you believe your friend will enjoy.

To have a false belief is to be mistaken about what the world is like, in some respect. As a result, you are more likely to achieve what you set out to do if you make decisions based on true beliefs rather than false ones. For that reason, whether your beliefs are true is a question of great importance.

However, there are many topics concerning which we cannot determine with absolute certainty what is true. That is, there is no procedure by which we can infallibly determine, for any given belief whatsoever, whether it is true or false. What, then, are we do to? One thing you could do is to simply believe whatever you happen to believe, without worrying about whether your beliefs are true or false. This does not necessarily mean your beliefs never change, for indeed you may come to believe things which you did not before, and you may cease to believe things which you once did. But at any given time, you believe whatever it is you believe, and that’s fine. Maybe you’ll change some of those beliefs later. That’s fine too.

The problem with this approach is that it seems too weak a response to the problem of false belief. Sure, in some cases you may cease to believe something because you are confronted with a very strong reason to think it is false, or you may come to believe something because you are confronted with a very strong reason to think that it is true. But, it could also happen that over time you end up with a greater and greater proportion of false beliefs and a smaller and smaller proportion of true beliefs. Presumably, this is not what you want.

This suggests a different response: working diligently to try to find what is likely to be true and what is likely to be false. In some sense, this is the job of scientists, mathematicians, philosophers, reporters, doctors, attorneys, judges, and so on. Of course, no individual could be expected to become an expert in all of these disciplines. Even if she could, science, mathematics, philosophy, and so on are still very much ongoing enterprises, and will probably continue to be so long as people exist.

The individual, however, can work diligently to try to find what is likely to be true and what is likely to be false given whatever evidence is available to her or him. This we can call

> **Critical thinking**: Seeking to find what is likely to be true and to avoid what is likely to be false given the evidence available to you.

The purpose of this book is to provide you with some tools to become a better critical thinker.

Starting Assumptions

In order to effectively engage in critical thinking, it is necessary to recognize two things:

1. *You are not certain that you will always believe what you have good reason to think is true, nor that you will never believe what you have good reason to think is false.*

2. *It is sometimes possible to determine whether a given belief is likely to be true given a certain body of evidence*.

If either of these claims were to be false, there would be no point in doing critical thinking.

You are not certain that you will always believe what you have good reason to think is true, nor that you will never believe what you have good reason to think is false. Suppose that you could be certain that you will always believe what you have good reason to believe, and that you will never believe what you have good reason not to believe. Were this the case, you would not need to engage in critical thinking. The purpose of critical thinking is to determine what is likely to be true or false given your evidence. If you always believe what you have good reason to believe, and never believe what you have good reason not to believe, then you already believe what is likely to be true given your evidence. Thus, there is no point in engaging in critical thinking.

Yet you cannot be certain that you will always believe what you have good reason to believe true, and that you will never believe what you have good reason to believe false. This is not impossible: it could just so happen by luck that you believe everything which you have very good reason to believe and disbelieve everything which you have very good reason not to believe. Still, this is wildly unlikely to happen, unless you engage in good critical thinking.

It is sometimes possible to determine whether a given belief is likely to be true given a certain body of evidence. Were this not the case, we would never be in a position to correctly say that a given belief is likely to be true. If that were so, critical thinking would be pointless. But this is not the case, as can be shown by example. Suppose that I place 10 balls in a container, one of which is red, the rest of which are blue. Furthermore, I do this is in front of you, where you can easily observe what I am doing. More, you were able to examine the container beforehand,

and are satisfied that it is an ordinary container. It is reasonable to believe that were you to draw a ball at random from the container, it would be blue.

Were either 1 or 2 to be false, critical thinking would be pointless. If 1 were false, the goal of critical thinking would have already been accomplished. If 2 were false, the goal of critical thinking would be impossible to achieve. As we have seen, however, it is reasonable to believe that both 1 and 2 are true.

Properly recognizing the truth of 1 means practicing intellectual humility: not only do you have no guarantee that all of your beliefs are correct, without critical thinking you cannot even be justifiably confident that all your beliefs are reasonable. Since it would very likely be in your benefit overall to have true rather than false beliefs, you have reason to seek out what is most likely to be true and to avoid what is likely to be false, if it is possible to do so. The fact that 2 is true means that it is sometimes possible to find what is most likely to be true and avoid what is most likely to be false. Thus, you have good reason to become a better critical thinker.

Work Problems 1.1

1. Think of a situation in which you are better off believing something false than something true.
   1.1. Are such situations likely to be common or uncommon?
   1.2. Why?
2. According to what is sometimes called *individual* or *subjective relativism*, whatever the individual believes to be correct, is correct. (A related view is *social* or *cultural relativism,* according to which whatever the members of a society believe to be correct, is correct.)

2.1. What would it mean for critical thinking, if individual relativism were correct? (Hint: consider our two starting assumptions for critical thinking).

2.2. Some people think individual relativism is self-defeating, meaning that if we think it is correct, in some cases we will be obliged to admit it is incorrect. Can you tell why?

Reasons for Belief: Justifications Verses Causes

It is in your benefit to practice good critical thinking. This means looking for what you have the best reasons to believe, including what you have the best reasons to believe you should and should not do. What this involves will be discussed in some detail throughout this book. At this juncture, however, it is useful to make a distinction which will help avoid a potential confusion.

When we talk about reasons for belief, we can mean different things. One thing we can mean is the causes of belief, that is, the events and processes which, given certain conditions, lead to a person adopting a certain belief. A second thing we can mean are the justifications for belief, that is, those pieces of evidence that might be offered for the belief being true.

In this book, our concern is primarily with reasons for belief in the second sense: justifications for belief. This distinction is an important one, because if we are not good critical thinkers we can be caused to have beliefs for which we do not have good justification. This can lead us to have false beliefs, as well as leaving us vulnerable to manipulation.

One of the most powerful factors in what we are caused to believe is our emotions: sometimes we believe things largely or entirely because of how they make us feel. People can use knowledge of our desires, hopes, and fears to manipulate us into believing and doing things which we may not have good justification to believe or do.

In an advertisement, a group of actors is seen having a fun time while consuming a certain brand of beverage. This is supposed to make you associate the brand of beverage with good feelings, making it more likely that you will purchase it. A demagogue gives a fiery speech, warning about a lurking danger, and then assures his audience that by all working together under his lead, they are sure to overcome that danger. This is supposed to make them more likely to vote for, or otherwise give the demagogue power. This is sometimes call **appeal to emotion**.

It is important not to misunderstand. The point is not that emotion is bad, or that trying to stir up emotion in people is always bad. If you are attuned to the way the world is, because you have true beliefs, then your emotions can be a valuable source of information: if something scares you, it might be because it is dangerous. Further, good speakers and writers will often try to elicit an emotional response in their audience. This is not necessarily a bad thing. Writings and speeches that fail to cause emotional reactions are often boring, and sometimes stirring up emotion can be useful to motivate people to do good things. The point is that, when someone is trying to get you to believe or do something, it is wise to consider what justification is being offered, if any.

Another significant factor in what we are caused to believe is the natural desire to feel that you are a part of certain groups, as well as the desire to feel that you are not a part of other groups. Your awareness that a belief is, or seems to be, common among members of a group you are or want to be a part of can, if you do not take efforts to avoid it, lead you to adopt the belief. Likewise, your awareness that a belief is, or seems to be, common among members of a group you do not want to be a part of can, if you do not take efforts to avoid it, lead you to reject the belief. This is another means through which people may try to manipulate you, and this is one version of what is sometimes called **appeal to popularity** or **appeal to the masses**.

Again, it is important not to misunderstand. The point is not that the beliefs held by members of a group are never good or bad. Sometimes they are. Nor is the point that you should never assume that if the members of a group belief or disbelieve something, that is a good reason to believe or disbelieve it. Sometimes it is. Indeed, if most experts on a given topic have a certain belief about that topic, that does constitute justification for accepting the belief.

The point is that you should seek out what justification there is for various beliefs, and that you should be conscientious about the various factors which might cause you to adopt beliefs for which there is inadequate justification. When you find yourself starting to wonder if such and such is the case, ask what justification you have for thinking it is or is not so. Especially whenever someone is trying to get you to believe or do something, step back and ask yourself whether this person is offering you *justification* for the belief.

Work Problems 1.2

1. Find an example of when emotion is used to try to get an audience to believe or do something.
    1.1. What is the audience being asked to do or believe?
    1.2. What justification, if any, is offered by why the audience should believe or do it?
1. Describe a situation in which using emotion to try to get an audience to believe or do something would likely be seen as a bad thing.

Chapter 2. Introducing Arguments

How do we decide what is likely to be true, given the available evidence? Well, we look to see what we have the most reason (in the sense of justification) to believe. Whenever someone offers a reason to believe something, that person is making an **argument**. The thing the person making an argument is trying to get you to believe (assuming that person is arguing sincerely), is called the **conclusion**. The reason or reasons offered for why you ought to believe the conclusion are called **premises**.

Let's look at two example arguments. First,

Fritz is a cat, and all cats are mammals. So, Fritz is a mammal.

Here, the premises are:

1. *Fritz is a cat*.
2. *All cats are mammals*.

The conclusion is:

Fritz is a mammal.

The second example goes like this,

Toby must like grapes, since he's a turtle and turtles usually like grapes.

Here, the premises are:

1. *He [Toby] is a turtle.*
2. *Turtles usually like grapes.*

The conclusion is:

Toby must like grapes.

Two things are worth noticing. The first is that, the conclusion comes at the end in the first argument, but at the beginning in the second. There is no general rule about where the conclusion must appear in an argument.

The second thing worth noticing is that the word “so” in the first argument indicates the conclusion, and the word “since” in the second indicates the premises. In fact, many words and phrases can function as **conclusion indicators** or **premise indicators**. Here are some examples:

Premise Indicators	**Conclusion Indicators**
Since	So
Given that	Therefore
Seeing that	Hence
Due to the fact that	Thus
For the reason that	It follows that
As evidenced by	In conclusion

Words and phrases like this can help you to identify arguments, and can be used to determine, for a given argument, what is the conclusion and what are the premises. Keep in mind, though, that many of these words and phrases have other uses. The word “so” in “Sally is so smart” is not introducing a conclusion.

To believe something means to think that it is true. This means that the real conclusion of an argument is always something that can be true or false. What sort of thing can be true or false? Well, **statements** can be true or false. Examples of statements include:

The sky is blue.

The sky is red.

Dinosaurs went extinct around 65 million years ago.

A fish named Phil runs a non-profit organization from my basement.

Tom is Jane’s uncle’s cousin’s friend.

Since the conclusion of an argument is what the person giving the argument wants you to believe (assuming that person is arguing sincerely), and to believe something is to believe it is true, the proper conclusion of an argument must be a statement.

A **statement** is a piece of language of the kind which can be either true or false. Not all sentences are statements. Questions, commands, and exclamations are all sentences (or can be expressed by sentences) which are neither true nor false. Here are some examples:

Questions	**Commands**	**Exclamations**
How are you?	Don’t talk to me!	Pizza, hooray!
Want to hang out later?	Be sure to study for the exam	Ouch!

Most competent language users can distinguish statements from other kinds of sentences with little effort.

What counts as a reason to believe that a statement is true? Well, possible states of affairs and events which, if they obtain or occur, mean that the statement is true or is likely to be true. You have reason to believe your house has been broken into, for instance, if the window is shattered, and your belongings are strewn about haphazardly rather than in their normal locations. You have reason to believe that a certain book is old if the pages are yellowed, the cover is worn, and the book has a musky smell.

In presenting a reason to believe something, someone might present some object as evidence. For instance, as reason to believe that Bill took Tonya’s sandwich, Fred might produce the remains of the partially eaten sandwich from Bill’s backpack. In giving an argument, someone presents reasons to believe a conclusion by *telling you* what those reasons are. You say that something is the case by making a statement. Thus, if I want to indicate that I am hungry, I say: “I am hungry.” As such, the premises of an argument are also statements, just like the conclusion.

We can now offer more robust definitions of the terms **conclusion**, **premise** and **argument**.

> **Conclusion**: The statement in an argument which a person sincerely making the argument is trying to get you to believe.
>
> **Premise**: A statement in an argument which, either by itself or together with other premises, is offered as a reason to believe the conclusion.
>
> **Argument**: A series of statements, one of which is the conclusion, the rest of which are premises, such that the premises are offered as providing reason to believe the conclusion.

Note importantly that the premises are *offered* as providing good reason to believe the conclusion. In fact, the premises might provide good reason to believe the conclusion, or they might not. A major component of good critical thinking is learning how to tell the difference.

Work Problems 2.1

1. For each of the following sentences, say whether it is a statement or a non-statement.
    1.1. I have a cat named Steven.
    1.2. Jessica enjoys hockey.
    1.3. Do your homework.
    1.4. Extraterrestrials secretly run the government.
    1.5. Stop interrupting me!

1.6. Brussel sprouts, yum!

1.7. Do you want to go to the party?

1.8. Jamal can help if you ask him.

1.9. Go ask Jamal if you want help.

2. Statements represent the ways the world could be, and they can be true or false. Are there any statements which a reasonable person might think are neither true nor false?

3. For each of the following arguments, identify the premise or premises and the conclusion.

   3.1. We know that some creatures lived and died off long before humans were around, because we have discovered the fossilize remains of plants and animals for which no human groups have records of living specimens.

   3.2. Even if the probability that any given star is orbited by a planet with life is incredibly small, it seems that there must be extraterrestrial life since there is an incredibly enormous number of stars that exist in the universe.

   3.3. If you want to do well for yourself you should be kind to others, because people repay favors.

   3.4. If you want to do well for yourself, you should be kind to others. I know you want to do well for yourself. So, you should be kind to others.

   3.5. If you can assign every item on one list to one item on a different list, so that every item on the second list has exactly one item assigned to it, there are the same number of items in each list. You can assign the number 2 to the number 1, the number 4 to the number 3, the number 6 to the number 5, and so on forever. There is the same number of odd and even numbers.

Suggested and Missing Premises and Conclusions

Consider the following passage:

Jim hasn't ever lied to you about something important before. Shouldn't you believe what he is telling you now?

Is this an argument? If it is, it must have a conclusion. What might this be? The answer clearly is not

Jim hasn't ever lied to you about something important before,

because no reason is offered for why the other person should believe this. It is simply assumed that the person to whom the argument is directed will agree. This leaves us with

Shouldn't you believe what he is telling you now?

But this is a question, not a statement! Does this mean that there is no argument here?

In fact, there is an argument here. It goes like this:

Jim hasn't ever lied to you about something important before.	(Premise)
You should believe what Jim is telling you now.	(Conclusion)

The tricky thing here is that the person making the argument does not explicitly state the conclusion. Rather, that person suggests the conclusion with a rhetorical question.

Also consider:

If you don't get dressed now, we'll be late. So, get dressed!

The real argument goes like this:

If you don't get dressed now, we'll be late.	(Premise)
So, you should get dressed.	(Conclusion)

In this case as well, the person making the argument does not explicitly state the conclusion. Instead of using a rhetorical question to suggest the conclusion, she or he uses a command.

Finally, consider:

Has Jim ever lied to you about something important before? Well, you should believe him now.

In this case, it is the premise which is not explicitly stated. The real argument goes like this:

Jim has never lied to you about something important before. (Premise)

You should believe Jim now. (Conclusion)

It is good to be aware that when people give arguments in the real world, they do not always state each premise and conclusion explicitly, but sometimes use a question or command to suggest it. When this occurs, it is important to identify the real premises and conclusion. You cannot properly evaluate an argument unless you can identify the premises and the conclusion.

In some cases, people give arguments in which a crucial premise is simply absent. Consider the following:

Since Fritz is a cat, Fritz must sleep a lot.

The only premise explicitly stated is

Fritz is a cat.

This premise by itself provides no reason whatsoever to believe the conclusion, since it does not tell us anything about Fritz or cats, other than that Fritz is one of them. The real argument is

Fritz is a cat. (Premise)

Cats sleep a lot. (Missing Premise)

Fritz must sleep a lot. (Conclusion)

Such arguments are called **enthymemes**. An enthymeme is:

An argument with an unstated premise or premises, or an unstated conclusion.

Typically, the missing premise of an enthymeme will be a piece of common belief or knowledge. Since the missing premise is commonly believed, the arguer takes the license of leaving it unstated.

For the purpose of evaluating arguments, it is important to identify missing premises. In some cases, these missing premises may be questionable. Consider the following:

Jill does not respect my religious beliefs, since she's an atheist.

Here is the argument in full:

Jill is an atheist.	(Premise)
Atheists are not respectful of others' religious beliefs.	(Missing Premise)
Jill must not respect my religious beliefs.	(Conclusion)

That missing premise is plausibly false. While there are surely some atheists who don't respect the religious beliefs of other people, it seems likely that many atheists do respect at least some religious beliefs. Making the missing premise explicit allowed us to identify a weakness in the argument.

Work Problems 2.2

For each of the following arguments, identify the missing premise or conclusion.

1. I don't trust John; he's a marketing major.
2. Monique Jones' last book received a prestigious award. It must be good.
3. Ralph is a dog, so he probably likes getting his belly rubbed.
4. You should be polite when you introduce yourself. She knows a lot of important people. You want to make a good impression, don't you?
5. Go talk to him! I bet he'd be willing to make amends if you apologize.

Types of Argument

Whenever a person sincerely offers an argument, that person is presenting reasons to think that the conclusion of the argument is true. Different kinds of reasons can constitute different kinds of support for a conclusion. We can classify arguments based on the kind of support the premises provide for the conclusion, if the argument is successful.

An argument is **deductive** just in case:

The premises either conclusively support the conclusion, or the argument fails.

An argument is **inductive** just in case:

The premises either make the conclusion likely, or the argument fails.

How you go about evaluating an argument depends in part on whether that argument is inductive or deductive, so it is good to be able to tell the difference.

It may be helpful to look at some examples. The first goes like this:

Fritz is a cat, and all cats are mammals. So, Fritz is a mammal.

Here it is again, with the premises and conclusion specified:

Fritz is a cat.	(Premise)
All cats are mammals.	(Premise)
Fritz is a mammal.	(Conclusion)

Notice that were the premises true, the conclusion would have to be true. That is what is meant by saying that the premises provide conclusive support for the conclusion. This argument is deductive.

Our second example is:

Toby must like grapes, since he's a turtle and turtles usually like grapes.

Here it is again, with the premises and the conclusion specified:

He [Toby] is a turtle.	(Premise)
Turtles usually like grapes.	(Premise)
Toby likes grapes.	(Conclusion)

Notice that even if the premises are true, it is possible for the conclusion to be false. Even if most turtles like grapes, Toby might not. Still, if it is true that most turtles like grapes, it is very likely that Toby likes grapes, since Toby is a turtle. This argument is inductive.

In this second argument, notice that the sentence expressing the conclusion contains the word "must." This might be taken to mean that the argument is deductive, because it suggests that if the premises are true the conclusion must be. However, the argument is clearly inductive, since if the premises are true the conclusion is likely. If you are not careful, words and phrases like "must," "it must be the case," "It is certainly true that," and others can lead you to think an inductive argument is really deductive. When deciding whether an argument is inductive or deductive, ask yourself what kind of support the premises provide the conclusion if the argument does not fail.

Our third example is:

Fritz is a mammal and a call cats are mammals. So, Fritz is a cat.

Here it is again, with the premises and the conclusion specified:

Fritz is a mammal.	(Premise)
All cats are mammals.	(Premise)
Fritz is a cat.	(Conclusion)

In this case, even if the premises are true, the conclusion could still be false. Fritz could be a dog, or a mouse, or a great many other things. Indeed, the premises do not even make it very likely that Fritz is a cat. What kind of support do we think a person making this argument would be trying to give for the conclusion? Well, there is nothing in here about likelihood. Recall that in our last example, one of the premises was

Turtles usually like grapes.

A fact which usually holds for turtles is likely to hold for a particular turtle, though not guaranteed. The word 'usually' in this premise introduces likelihood into the argument. Since there is nothing introducing likelihood in this third argument, we should conclude that it is deductive.

Our fourth argument goes like this:

Toby must like grapes, since he's a turtle and several turtles like grapes.

Here it is again, with the premises and the conclusion specified:

He [Toby] is a turtle.	(Premise)
Several f turtles like grapes.	(Premise)
Toby likes grapes.	(Conclusion)

The fact that a lot of turtles likes grapes provides us with some reason to think that Toby the turtle likes grapes, so the argument is inductive. However, the fact that several turtles like grapes doesn't give us very much reason to think that Toby likes grapes. How many is a lot? It is not clear.

Here is another argument:

So far, every time I've flipped the switch on my wall up, the light came on. So, the next time I flip the switch on my wall up, the light will come on.

And again, with the premise and conclusion specified:

So far, every time I've flipped the switch on my wall up, the light came on. (Premise)

The next time I flip the switch on my wall up, the light will come on. (Conclusion)

Is this argument deductive or inductive?

Clearly even if the premise is true, the conclusion might still be false: maybe the next time I flip the switch up the power is out at my house. It also contains no explicit mention of probability. However, the argument involves probability *implicitly*. Suppose, as seems likely, that I have flipped the switch up many, many times, and on every such occasion the light has turned on. It would be *unlikely* for these events to consistently occur together just as a matter of change. Thus, it is *likely* that they are connected. The argument is inductive.

Yet another argument is:

Cutting military spending would help reduce the deficit. We should cut military spending, since it is good to reduce the deficit.

Here is the argument with the premises and the conclusion specified:

Cutting military spending would help reduce the deficit.	(Premise)
It is good to reduce the deficit.	(Premise)
We should cut military spending.	(Conclusion)

The second premise and the conclusion are both normative statements. A **Normative Statement** is:

A statement about what we have reason to do or not to do.

People sometimes get confused whether an argument with a normative conclusion is deductive or inductive.

To decide whether an argument with a normative conclusion is deductive or inductive, consider whether, if the reasons offered to do or not do the act were to be accepted, would it be *conclusive* or *inconclusive*? If we accept that it is good to reduce the deficit, then that gives us some reason to cut military spending, assuming doing so would indeed reduce the deficit. Even so, this reason is not conclusive, as there may be good reasons not to cut military spending which might outweigh the benefits of reducing the deficit.

Here is our final argument:

It is always wrong to torture someone merely for fun. Robert has no reason to torture Sam except that he thinks it will be fun. *Robert should not torture Sam*.

And again, with the premises and the conclusion specified:

It is always wrong to torture someone merely for fun.	(Premise)
Robert has no reason to torture Sam except that he thinks it will be fun.	(Premise)
Robert should not torture Sam.	(Conclusion)

This argument also has a normative conclusion. Is it deductive or inductive?

Well, is the reason offered not to torture Sam, if we accept it, conclusive or inconclusive? According to the first premise, torturing someone merely for fun is *always* wrong. It is part of the meaning of the word "wrong" that one should never do what is wrong. (By contrast, it is sometimes permissible to do something with bad effects, if these are outweighed by the good effects). Even so, an act which is *typically* wrong may be permissible in some exceptional circumstances. If something is always wrong, however, then there are no such exceptions. Hence, this argument is deductive.

Work Problems 2.3

For each of the following arguments, indicate whether it is inductive or deductive.

1. We know that some creatures lived and died off long before humans were around, because we have discovered the fossilize remains of plants and animals for which no human groups have records of living specimens.
2. Even if the probability that any given star is orbited by a planet with life is incredibly small, it seems that there must be extraterrestrial life due to the incredibly enormous number of stars that exist in the universe.
3. If you want to do well for yourself you should be kind to others, because people repay favors.
4. If you want to do well for yourself, you should be kind to others. I know you want to do well for yourself. So, you should be kind to others.
5. If you can assign every item on one list to one item on a different list, so that every item on the second list has exactly one item assigned to it, there are the same number of items in each list. You can assign the number 2 to the number 1, the number 4 to the number 3, the number 6 to the number 5, and so on forever. There is the same number of odd and even numbers.
6. There are only fifteen miles left until we reach Cityville. That's what the sign says

Validity, Soundness, Strength, and Cogency

Deductive arguments can be divided into those which are valid and those which are invalid. An argument is **valid** just in case:

It cannot both be the case that the premises are all true and the conclusion false.

For a valid argument, in other words, if the premises are all true, the conclusion must be true. If a given argument is valid, the premises are said to **entail** the conclusion. A deductive argument which is not valid is **invalid**. Some valid arguments are sound. An argument is **sound** just in case:

It is valid and has only true premises.

A deductive argument which is not sound, either because it is invalid or because it a has at least one false premise, is **unsound**.

If an argument is sound, the conclusion must be true. After all, a sound argument is valid, and in a valid argument, if the premises are all true, the conclusion must be true as well. But in a sound argument the premises are all true as well. So, the conclusion must be true.

Of course, an unsound argument could have a true conclusion. Consider:

Fritz is a mammal and a call cats are mammals. So, Fritz is a cat.

This argument is invalid, since there is a way for the premises to all be true but the conclusion false. Since it is invalid, it is also unsound. However, it could be that the conclusion is true: Fritz may, in fact, be a cat. Or consider,

Toby is a mammal and all mammals are turtles. So, Toby is a turtle.

This argument is unsound since one of the premises,

All mammals are turtles,

is false. Still, the conclusion could yet be true: Toby may be a turtle.

Finally, a valid argument could have a false conclusion. Consider,

My best friend is a talking fish and all fish have feathers, so my best friend has feathers.

This argument is valid, since were the premises true, the conclusion would have to be true as well. But the premises and the conclusion are all false.

Only a sound argument must have a true conclusion.

Inductive arguments can be divided into those which are strong and those which are weak. An argument is **strong** just in case:

It is unlikely to both be the case that the premises are all true and the conclusion is false.

For a strong argument, in other words, if the premises are all true, the conclusion is likely to be true. An inductive argument which is not strong is **weak**. Some inductive arguments are cogent. An argument is **cogent** just in case:

It is strong and has only true premises.

The conclusion of a cogent argument is likely to be true, but not certain.

How likely must the premises make the conclusion in order for an inductive argument to be strong? There is no established answer to this, but in general the more likely a conclusion is given the premises, the stronger than argument. This reveals an important difference between inductive and deductive arguments. In inductive arguments, the amount of support the premises provide for the conclusion admits of degree: it can be more or less. In a deductive argument, the support the premises provide for the conclusion is all or nothing.

There is another difference related to this one. A valid argument will remain valid no matter what additional premises are added to it. Consider the following argument, which earlier we said was valid:

Fritz is a cat and a call cats are mammals. So, Fritz is a mammal.

This argument is valid because there is no way for the premises to be true and the conclusion false. Is there any premise we could add which would make it so that

Fritz is a cat

and

All cats are mammals

are both true, but

Fritz is a mammal

is false? Clearly not, because any additional premise would not change whether or not Fritz is a cat given that Fritz is a mammal and all mammals are cats.

By contrast, by adding additional premises we can make a strong inductive argument weak. Consider the following argument:

> *90% of students who received an A in Professor Bongle's class were women, so this randomly selected student who received an A in Professor Bongle's class is a woman*.

Clearly, if the premise

> *90% of students who received an A in Professor Bongle's class were women*

is true, the conclusion,

> *This randomly selected student who received an A in Professor Bongle's class is a woman*,

is very likely to be true, and so the argument is strong. But suppose we add the following additional premise:

> *The randomly selected student is named John*.

Since it is highly unlikely that a woman would be named John, the argument is now weak.

Work Problems 2.4

Consider the following argument:

> *All dogs like crossword puzzles.*
>
> *Not all dogs like crossword puzzles.*
>
> *Therefore, some cats like French cuisine*.

1. Is this argument valid? (Many beginning students have trouble with this question. Think of the definition of validity very carefully when trying to answer it)

2. Is this argument sound?

The Role of Arguments in Critical Thinking

Earlier we said that critical thinking is seeking out what is likely to be true and to avoid what is likely to be false given the evidence available to you. We might equivalently define critical thinking as:

Seeking out what you have the most justification to believe.

Often times, people will try to present you with reasons to believe something. In doing so, they will be presenting you with arguments. Critical thinking involves seeking to discover whether those arguments are any good. On the other hand, at any given time you are aware of, or you are in a position to come aware of, many objects, states of affairs, or events which constitute potential evidence that some statements are true. Critical thinking involves seeking to discover what beliefs your evidence best supports. Reasoning from such evidence will be done through arguments. Thus, the evaluation of arguments plays a crucial role in critical thinking.

In evaluating an argument, we want to know two different things. First,

Do the premises provide support for the conclusion?

Second,

Are the premises true, or do we have good reason to believe they are true?

The answers to these questions are independent. We can try to find out if the premises are true, or if we have good reason to believe they are true, without first seeing if they support the conclusion. On the other hand, we can try to decide whether the premises provide support for the conclusion of an argument even if we have no idea whether the premises are true or not. We can, after all, imagine that the premises are true and consider whether the conclusion must be true or is likely to be true in that imagined case.

If you come across an argument such that the answer to both questions is no, then that argument provides you with no good reason to believe the conclusion. If you come across an argument such that the answer to both questions is yes, then that argument provides you with good reason to believe the conclusion.

Part II. Deductive Arguments

Chapter 3. Introducing Categorical Logic

An argument is deductive, recall, if the premises either provide conclusive support for the conclusion or the argument fails. If the premises of the argument do provide conclusive support for the conclusion, in the sense that it cannot both be the case that the premises are all true and the conclusion is false, then the argument is valid. Otherwise, it is invalid.

We would like to be able to tell whether a given deductive argument is valid or invalid. For, if we know that an argument is invalid, we know that the premises do not provide any good reason to believe the conclusion. On the other hand, if we know the argument is valid, we know that if the premises are true the argument is sound, meaning the conclusion must be true.

In some cases, it can be pretty easy to tell whether an argument is valid or invalid just by thinking about it for a few moments. In other cases, however, it might be harder to tell, and it is easy to make mistakes if due care is not taken. If possible, we would like some procedure for identifying valid arguments.

In fact, at least for certain kinds of deductive arguments, there are procedures for deciding if those arguments are valid. To learn such a procedure requires learning about a logic. A logic consists of

- A set of rules specifying the kinds of statement dealt with by the logic
- A set of procedures by which one may determine whether a given collection of statements entails some other statement.

The latter constitutes a **proof system** for the logic.

The topic of this chapter and the next is **categorical logic**. It is called “categorical logic” because the statements it deals with concern categories of things: birds, trees, people, red things, square things, and so on.

Translating into Categorical Form

Categorical logic only deals with arguments composed out of certain types of statements. Each has the following basic structure:

Quantifier *subject term* *copula* *predicate term.*

Here are some examples:

All humans are animals.

Some foxes are clever things.

No humans are foxes.

Some clever things are not humans.

The quantifier is “All” “Some” or “No”. The copula is “are” or “are not.” The subject terms in the examples above are “humans,” “foxes,” “humans,” and “clever things,” whereas the predicate terms are “animals,” “clever things,” “foxes,” and “humans.” A term in categorical logic is always either a noun or a noun-phrase, which is a phrase which functions as a noun. A statement with this basic structure is said to be in **categorical form**.

In categorical logic, there are four basic kinds of statement, or **statement forms**. A **Statement form** is the basic structure of a statement independent of what it is about. These are traditionally represented by the letters A, E, I, and O:

A. All P are Q. E. No P are Q.

I. Some P are Q. O. Some P are not Q.

The letters P and Q are variables which may be replaced with terms, resulting in a statement in categorical form. What a statement about is called its **content**.

Each statement form has a **quantity** and a **quality**. The quantity is either **universal** or **particular**. The quality is either **affirmative** or **negative**. The quantity and quality of each of the four kinds of statement is given by this chart:

	Affirmative	**Negative**
Universal	A. All P are Q	E. No P are Q
Particular	I. Some P are Q	O. Some P are not Q

A and I statements have an **Affirmative** quality, whereas E and O statements have a **Negative** quality. A and E statements have a **Universal** quantity, whereas I and O statements have a **Particular** quantity. Thus, A statements are **Universal Affirmative**, E statements are **Universal Negative**, I statements are **Particular Affirmative**, and O statements are **Particular Negative**.

In order to evaluate an argument in a natural language like English using categorical logic, it is sometimes first necessary to put the statements contained in that argument into categorical form. Consider the statement

All birds have feathers.

This statement is clearly about all of something: all birds. So, it is an A statement. To put it into categorical form, we must rewrite it with the form *All P are Q*. But

All birds are have feathers

is obviously ungrammatical. So, we have to replace *have feathers* with a noun or noun-phrase. Thus, we have:

All birds are things with feathers.

Similarly, the statement

Whatever has feathers lays eggs

becomes

All things with features are things that lay eggs.

Putting statements into categorical form can be challenging at first, but it becomes fairly easy with practice.

Here is a hint for making sure the subject and predicate terms are in the correct format: try switching them and see if the result is grammatically correct. For instance, take the statement

All cats are furry.

Switching the subject and predicate terms gives us:

All furry are cats,

which is ungrammatical. So, “All cats are furry” is not in categorial form. On the other hand, if you switch the subject and predicate terms of

All cats are furry things,

what you get is

All furry things are cats.

The latter statement may be false, but it is grammatically fine. So, “All cats are furry things” is in categorical form.

Sometimes an English statement will not contain any explicit indication of the quantity. For example, consider

Kids like chocolate.

All kids or some kids? When putting such statements into categorical form, you should use whichever reading seems most plausible in the context. In this case, we know that not all kids like chocolate, so it is probably good to go with

Some kids are things that like chocolate.

Notice, that we replaced "like chocolate" with the copula "are" and the noun-phrase "things that like chocolate" when putting the statement into categorical form.

One kind of statement deserves special attention: **singular statements**. A singular statement is a statement about one particular thing, such as "Toby is a turtle." To put a singular statement into categorical form, we treat it as a statement about all things identical to that particular thing. Thus,

Toby is a turtle

becomes

All things identical to Toby are turtles.

Similarly,

Toby likes grapes

becomes

All things identical to Toby are things that like grapes.

Obviously, these are both A statements.

Now consider a singular statement like *Jill is not temperamental*. Here is the statement rewritten in categorical form:

All things identical to Jill are non-temperamental.

Is this an A statement? Well, yes, technically. There is no rule which says *non-temperamental* cannot be a predicate term. (Importantly, note that *non-temperamental* is not the same as *not* *temperamental*. The *non-* in *non-temperamental* is a part of the term, whereas in *not temperamental*, the term is just *temperamental* with *not* functioning to indicate the quality.) Although this is an A statement, it is clearly different from other A statements. This difference will be discussed later, in the next chapter. For now, it is enough to know to translate such singular terms into categorical form.

Work Problems 3.1

Rewrite each of the following statements in categorical form and indicate the statement type.

1. Dogs are mammals.
2. Cats like being petted.
3. Birds don’t have fur.
4. Some people don’t care for rock and roll.
5. Some hip hop fans like country music.

Argument Forms

Once we identify the types of statement in an argument, we can isolate the **argument form**. The argument form is the basic structure of the argument, independent of its content. In the case of categorical logic, that means the argument form ignores the meaning of the subject and predicate terms. To find the argument form of an argument, we simply replace the subject and predicate terms of the argument with capital letters, making sure that a given term is replaced with the same letter every time it appears in the argument. These capital letters are variables standing for terms.

For instance, take the argument:

All birds are things with feathers.

All things with feathers are things which lay eggs.

All birds are things which lay eggs.

We can replace “birds” with “*P*”, “things with feathers” with “*Q*” and “things which lay eggs” with “*R*”. This gives us:

All P are Q

All Q are R

All P are R

Or consider this argument:

All birds lay eggs.

Some things which lay eggs are lizards.

Some birds are lizards.

Replacing “birds” with “*P*”, “things which lay eggs” with “*Q*” and “lizards” with “*R*” gets us:

All P are Q

Some Q are R.

Some P are R.

Notice that these two arguments have the same form.

Every argument form is a pattern for a potentially infinite number of arguments. To get an argument from an argument form, simply replace the variables with the appropriate kind of content. In categorical logic, variables like “*P*”, “*Q*”, and “*R*” stand for nouns and noun-phrases; so, we can get an argument from an argument form by replacing these variables with nouns or noun-phrases.

Suppose we have the following argument form:

All P are Q.

All Q are R

All P are R

To get an argument, we could replace "*P*" with "dragons," "*Q*" with "things that breathe fire," and "*R*" with "things which terrorize the countryside," yielding:

All dragons are things that breathe fire.

All things that breathe fire are things which terrorize the countryside.

All dragons are things that terrorize the countryside.

In this case "dragons," "things that breathe fire," and "things which terrorize the countryside" are the **content** of the argument. When doing logic, we often look at the form of an argument independent from its content.

Earlier, in chapter 2, we explained what it means for an argument to be valid: it cannot both be the case that the premises are all true and the conclusion is false. The word "valid" can apply to argument forms as well. An argument form is valid just in case

Every argument with that argument form is a valid argument.

An argument form is not valid just in case

At least one argument with that argument form is not a valid argument.

The ability to identify valid arguments is crucial for good critical thinking, since a valid argument with all true premises (that is, a sound argument) must have a true conclusion. If you know that a certain argument form is valid, and you are able to correctly identify when an argument has that form, you are in a position to identify a potentially infinite number of valid arguments. Part of the utility of a logic is that it allows us to determine that certain argument forms are valid.

Work Problems 3.2

1 For each of the following arguments, identify the argument form.’

1.1 No humans are foxes.

No foxes are humans.

1.2 All humans are animals

Some humans are clever things.

Some animals are clever things.

1.3 Some humans are clever things

Some foxes are clever things.

Some humans are foxes.

1.4 All things identical to Toby are things that like grapes.

All things that like grapes are friends of mine.

All things identical to Toby are friends of mine.

2. For each of the following argument forms, write an argument with that form.

2.1. All A are B

Some B are C

Some A are C

2.2. No S are F

Some S are G

No F are G

2.3. Some A are not E

Some E are S

Some A are not S

2.4. No A are B

No B are C

No A are C

The Counterexample Method

Shortly we will explain a technique for determining that a valid argument in categorical logic is, in fact, valid. First, however, we will discuss a method for showing that a given argument form is not valid. This is called the **counterexample method**.

Recall that an argument form is not valid just in case there is at least one invalid argument with that form. Following the counterexample method, we identify the form of an argument and show that some argument with that form is obviously invalid.

First, we take an argument. For example,

Some teachers are smart people. (Premise)

Some smart people are lawyers. (Premise)

Some teachers are lawyers. (Premise)

Next, we identify the argument form. In this case, we have three terms: *teachers*, *smart people*, and *lawyers*. Since we have three terms, we need three term letters. Let us pick "*P*" for "teachers", "*Q*" for "*smart people*", and "*R*" for "*lawyers*". Substituting throughout gets us:

Some P are Q

Some Q are R

Some P are R.

Next, we see if we can come up with an argument with this form which is clearly invalid. That is, we try to come up with an argument having this form where the premises are obviously true, and the conclusion is obviously false.

How do we do this? Well, it is often useful to begin with the conclusion. So, we want a statement of the form "*Some P are R*" which is obviously false. For example,

Some lizards are birds.

Each variable has to have the same meaning throughout the argument. So, whatever term we substitute for "*P*" must be substituted wherever "*P*" occurs, and similarly with "*Q*" and "*R*". This gives us:

Some lizards are Q.	(Premise)
Some Q are birds.	(Premise)
Some lizards are birds.	(Conclusion)

Now we just need to substitute something for "*Q*" to make both premises obviously true, like so:

Some lizards are things with tails.

Some things with tails are birds.

Some lizards are birds.

Since the premises of this argument are obviously true and the conclusion is obviously false, we know the argument is invalid. But that means the argument form is not a valid argument form.

For comparison, consider a different argument.

All tigers are mammals.	(Premise)
All mammals are warm-blooded things.	(Premise)
All tigers are warm-blooded things.	(Conclusion)

What happens if we try to use the counterexample method on this argument?

Well, here is the argument form:

All P are Q.

All Q are R.

All P are R.

First, we make the conclusion something obviously false. For example,

All people are lizards.

Making the appropriate substitutions throughout the argument gives us:

All people are Q.

All Q are lizards.

All people are lizards.

Is there anything we can substitute for *Q* which would make both premises true? In fact, anything we substitute for Q will make at least one of these premises false. This is not surprising, since the argument form is valid.

Work Problems 3.3

For both of the following arguments, identify the argument form and use the counterexample method to show that the argument form is not valid.

1. Some humans are clever things.

   Some foxes are clever things.

   Some humans are foxes.

2. No birds are lizards.

   No lizards are things that do magic.

   No birds are things that do magic.

Validity, Argument Forms, and Enthymemes

Any argument with a valid argument form is valid. Also, every valid argument has a valid form.

Consider the following argument:

Fred is a triangle.	(Premise)
Fred has three sides.	(Conclusion)

Is this argument valid?

It is pretty obvious that this argument does not have a valid form. Here is a counterexample:

Fred is a fish.	(Premise)
Fred has feathers.	(Conclusion)

It is easy to imagine the premise being true while the conclusion is false. However, you are probably somewhat inclined to judge the original argument to be valid. After all, if Fred is a triangle, then Fred must have three sides, right? In making this judgment, you are implicitly relying on the fact that a triangle is a three-sided figure.

Consider this argument:

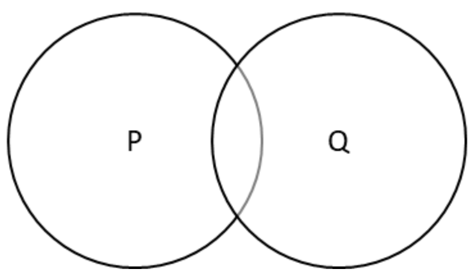
Fred is a triangle.	(Premise)
All triangles have three sides.	(Premise)
Fred has three sides.	(Conclusion)

This argument is valid. The original version of the argument is not valid as stated but can be considered valid if interpreted as an enthymeme.

Proving Validity with Venn Diagrams

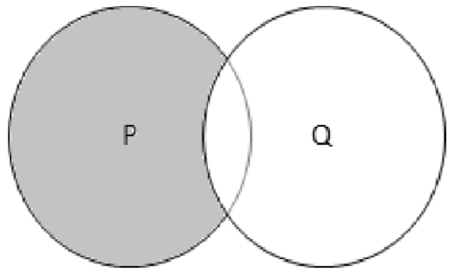
Recall that an argument is valid just in case it cannot be the case both that the premises are all true and the conclusion false. This means that, one way of showing that a valid argument is valid is to represent the premises as being true, and then determine if by doing so, we have already represented the conclusion as being true. If in the process of representing the premises of an argument as being true we also represent the conclusion as being true, then the argument is valid. However, if after representing the premises as being true, we could still consistently represent the conclusion as being false, then the argument is not valid.

We can use Venn diagrams to represent the premises of an arguments in categorical logic as being true, and in that way test those arguments for validity. A Venn diagram is a image consisting of some number of partially overlapping circles. When testing an argument in categorical logic for validity, start with a Venn diagram that has a number of circles equal to the number of terms in the premise or premises of the argument under evaluation. If the premise or premises of the argument contain only two terms, the Venn diagram will look like this:



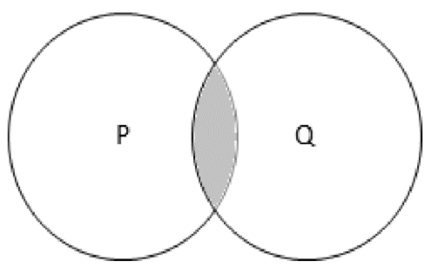
Each circle represents one term. The circle marked "P" represents everything which falls under the category designated by the variable "P", whatever that may be. Likewise, the circle marked "Q" represents everything which falls under the category designated by the variable "Q". The middle section, where the circles overlap, represents everything which falls under both the categories P and Q.

After making the Venn diagram, our next task is to use the diagram to represent the premise or premises of the argument as being true. Each of the four basic statement types in categorical logic (A, E, I, and O), can be represented on a Venn diagram with two circles. A statements, which have the form “All P are Q,” we represent as being true by shading in the part of the P circle which does not overlap with the Q circle. Thus:



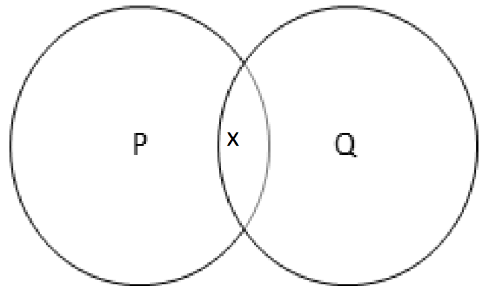
A shaded region indicates that the region is empty. The part of the P circle which does not overlap the Q circle represents everything which falls under the category P, but does not fall under the category Q. The statement “All P are Q” means that there is nothing which falls under the category P but fails to fall under the category Q. Thus, we shade to show that region of the diagram is empty.

E statements, having the form “No P are Q,” we represent as being true by shading the section where the P and Q circles overlap. So:



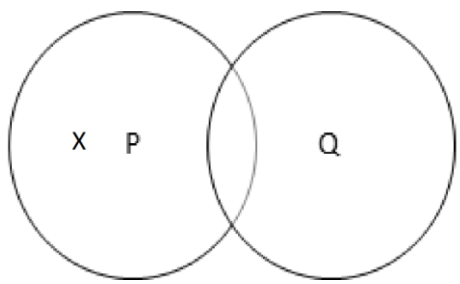
As before, the middle region is shaded to indicate that it is empty: nothing is both a P and a Q.

I statements, of the form “Some P are Q,” we represent as being true by placing an X in the middle region where the P and Q circles overlap. Hence:



An X in a region represents that something exists which falls under the category or categories of that region. Thus, an X in the middle region where the P and Q circles overlap represents that there exists something which falls under both category P and category Q.

O statements, having the form "Some P are not Q," we represent as being true by placing an X in the part of the P circle which does not overlap with the Q circle. Thus:



Since the section of the P circle which does not overlap the Q circle represents whatever falls under the P category but does not fall under the Q category, this represents the existence of something which is a P but is not a Q.

In each of these diagrams, we have let the left circle represent whatever falls under the P category, and we have let the right circle represent whatever falls under the Q category. Since P and Q can stand for any category at all, it would make no difference if we reversed them, letting the right circle represent whatever falls within the P category and the left circle represent whatever falls within the Q category.

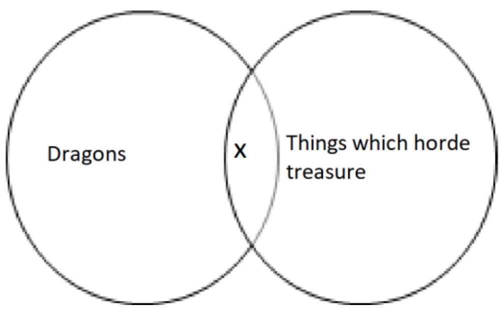
Recall that an argument is valid just in case it cannot both be the case that the premises are all true and the conclusion is false. Consider the argument

Some dragons are things which horde treasure (Premise)

Some things which horde treasure are dragons (Conclusion)

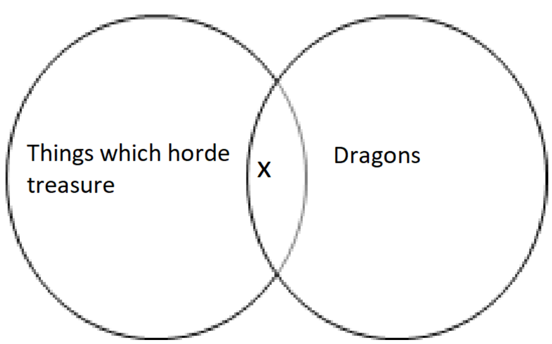
To test whether this argument is valid, we see whether it is possible to represent the premise as being true but not represent the conclusion as being true. If this is possible, then there is a way for the premise to be true and the conclusion false, and so the argument is invalid. If it is not possible, the argument is valid.

The premise of this argument, "Some dragons are things which horde treasure" contains two terms, and it is an I statement. So, we take the Venn diagram for I statements, and substitute "dragons" for "P" and "Things which horde treasure" for "Q," which gives us:

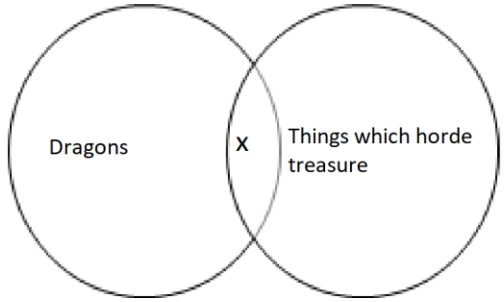


This represents the premise as being true.

Now, how would we represent the conclusion being true? The conclusion, recall, is "Some things which horde treasure are dragons." Clearly, this is an I statement. Substituting "Things which horde treasure" for "P" and "dragons" for "Q" in our Venn diagram for I statements gives us:



Recall that it does not matter which circle is on the right and which is on the left. This being the case, this Venn diagram represents exactly the same thing being true as this one:



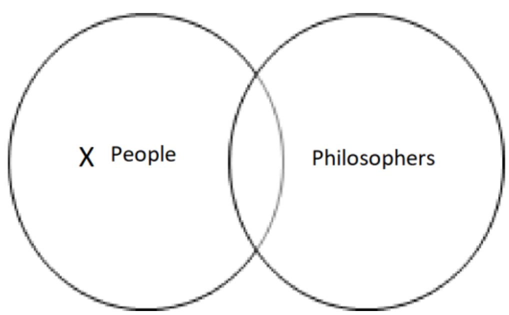
Thus, in representing the premise of our argument as true on a Venn diagram, we automatically represented our conclusion as being true as well. Our argument is valid.

Compare a different argument:

Some people are not philosophers. (Premise)

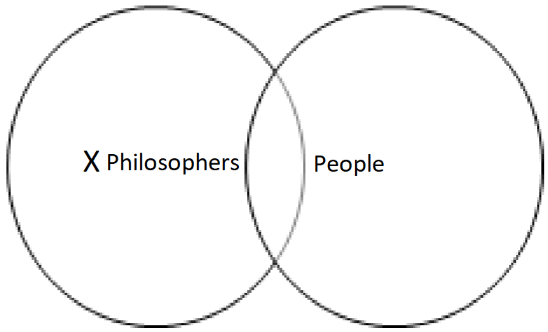
Some philosophers are not people. (Conclusion)

As before, to test whether this argument is valid, we see whether it is possible to represent the premise as true without representing the conclusion as true. The premise, "Some people are not philosophers," contains two terms and it is an O statement. Substituting "People" for "P" and "Philosophers" for "Q" in our Venn diagram for O statements gives us:



This represents the premise as being true.

Now, how would we represent the conclusion as being true? Recall the conclusion is "Some philosophers are not people." This is also an O statement. Substituting "Philosophers" for "P" and "People" for "Q" gives us:



This is clearly not the same as the Venn diagram for the premise, since the "X" is in a different section of the two diagrams. This means that we can represent the premise as being true without representing the conclusion as being true. Thus, there is a way for the premise to be true without the conclusion being true. Hence, the argument is invalid.

We have now used Venn diagrams to test two arguments for validity. One turned out valid, and the other invalid. The first argument had the following argument form:

Some P are Q. (Premise)

Some Q are P. (Conclusion)

Notice that, in showing that our argument was valid, we did not consider the meaning of either "Dragons" or "Things which horde treasure." This means that, had we replaced "P" and "Q" throughout the argument form with anything at all, the resulting argument would still be valid. Thus, this is a **valid argument form**.

The second argument had the following argument form:

Some P are not Q. (Premise)

Some Q are not P. (Conclusion)

Notice that, in showing that this argument was invalid, we did not consider the meaning of either "Philosophers" or "People." It is obvious that we could have replaced "P" and "Q" throughout the argument with different terms and still get an invalid argument. This is not a valid argument form.

Work Problems 3.4

For each of the following arguments, use a Venn diagram to show that it is (logically) valid, or use the counterexample method to show it is invalid.

1. No humans are foxes.

   No foxes are humans.

2. Some cats are mammals.

   Some mammals are cats

3. Some birds are not friendly things.

   Some friendly things are not birds.

4. All fish are things that live in water.

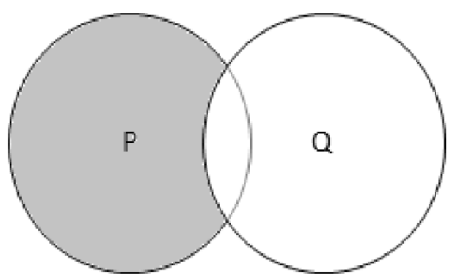
   All things that live in water are fish.

Chapter 4. Categorical Logic Continued

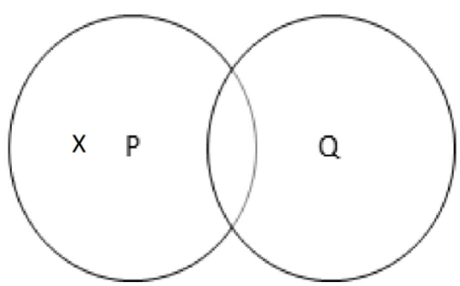
In the previous chapter we introduced categorical logic and discussed how to evaluate simple arguments for validity. In this chapter we continue our discussion of categorical logic. First, we take a closer look at the logical relations between our four basic statements forms. Then we consider how to evaluate arguments with three terms. Finally, we examine various ways of altering our four basic statement forms.

Squares of Opposition

Let us take a closer look at the logical relations between our four statement forms: A, E, I, and O. Begin with A and O. A statements, of the form "All P are Q" have the following Venn diagram:



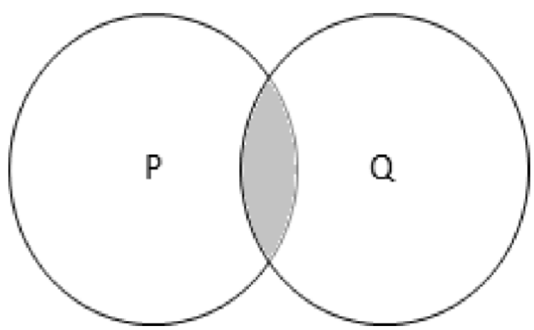
The Venn diagram for O statements, which have the form "Some P are not Q," looks like this:



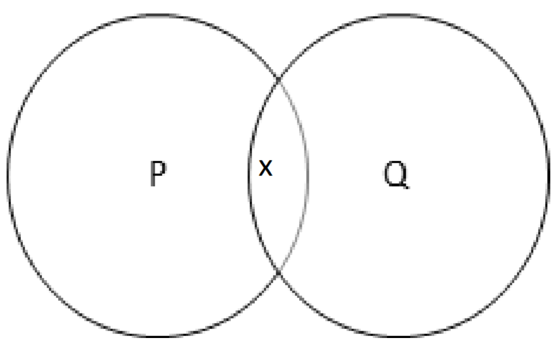
Remember that a shaded area means there is nothing which falls under that category or categories represented by that area, and that an X represents the existence of something which falls under that category. This means that if you shade in an area, you cannot also put an X in that area. Similarly, if you put an X in an area, you cannot also shade it.

Given this, we see that if "P" and "Q" have the same meaning in both statements, then "All P are Q" and "Some P are not Q" cannot both be true, nor can both be false. Clearly, if "All P are Q" is true, then "Some P are not Q" is false, and if "Some P are not Q" is true, then "All P are Q" is false. But "All P are Q" can only be false if there is some P which is not a Q, and so "Some P are not Q" is true. Likewise, "Some P are not Q" is false just in case there is not a single P which fails to be a Q, and so "All P are Q" is true. So, it must be the case that exactly one of "All P are Q" and "Some P are not Q" is true, and the other is false. For this reason, A and O statements are said to be **contradictories.**

E and I statements are also contradictories. The Venn diagram for E statements, of the form "No P are Q," looks like this:

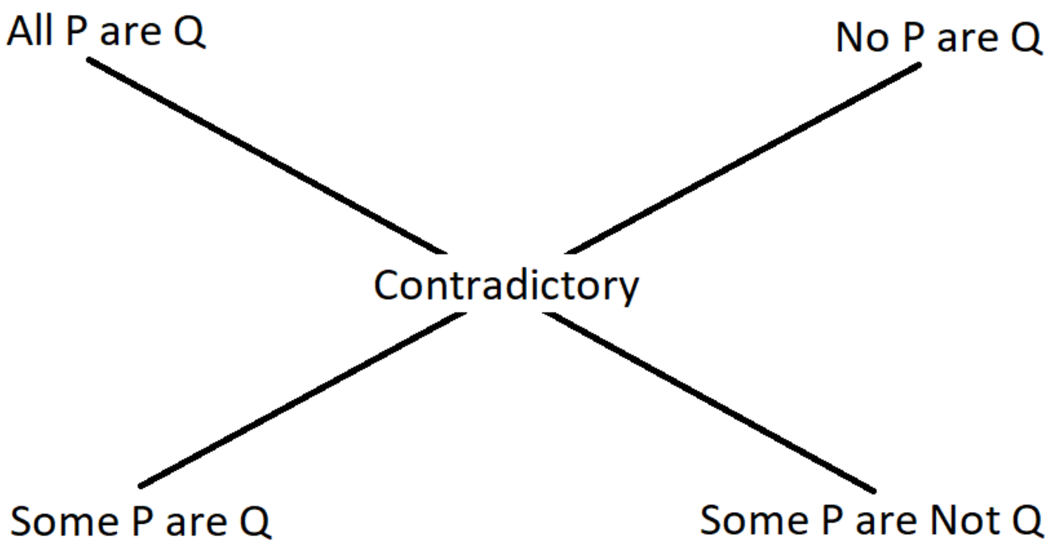


The Venn diagram for I statements, of the form "Some P are Q," looks like this:



Since a shaded region is understood to mean that nothing falls under that category, and an X in a region means that there is something which falls under that category, it is obvious from the Venn diagrams that if *P* and *Q* have the same meaning in both statements, exactly one of “No P are Q” and “Some P are Q” will be true and the other false.

We can represent this with the following diagram:



This is the **modern square of opposition**. Statement forms on opposite corners are contradictories: it must be the case that one is true and the other is false.

This is called the modern square of opposition in contrast with the **classic square of opposition**. The fundamental difference between the classic and the modern squares is this: in the classic square, it is assumed that there exists something which falls under the category P. This is not assumed in the modern square. Thus, we say that the classic square of opposition is **existentially committing**.

What is the significance of treating categorical statements as existentially committing? Consider an A statement, like “All giraffes are herbivores.” Does this statement validly entail the I statement “Some giraffes are herbivores” or not? If the original statement is existentially committing, then we assume there is something which falls under the category *giraffe*. That is, we assume there exists at least one giraffe. If that is the case, and the statement “All giraffes are herbivores” is true, then *those* giraffes must be herbivores. So, if “All giraffes are herbivores” is true”, “Some giraffes are herbivores” will also be true. Hence, if we assume existential commitment, an A statement validly entails an I statement with the same subject and predicate terms.

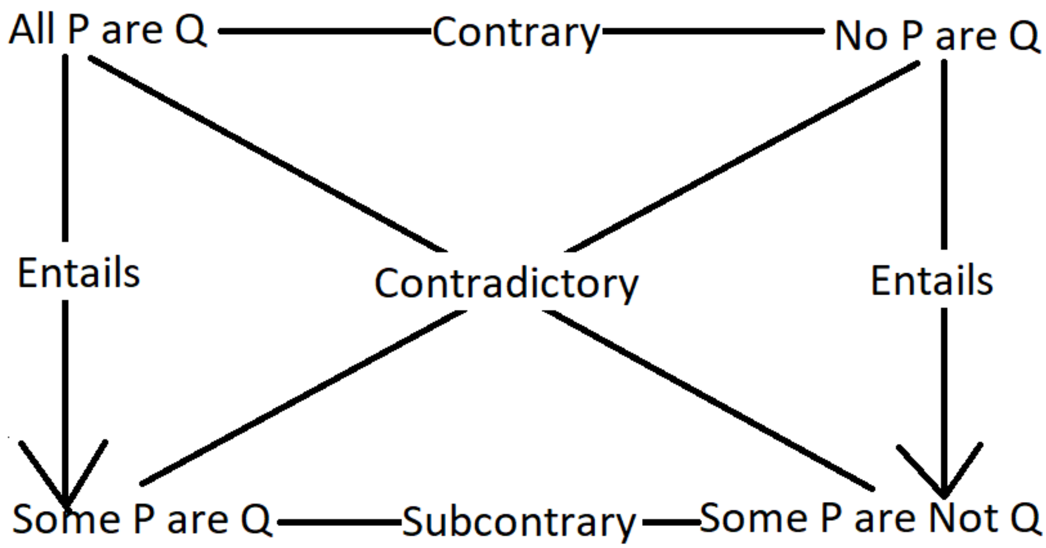
Compare the case in which we do not assume existential commitment. The statement “All giraffes are herbivores” could be true even if there are no giraffes. If this seems strange, consider a different A statement: “All unicorns are things with horns.” It seems like this could be true even if unicorns do not exist. If unicorns do not exist, then the I statement “Some unicorns are things with horns” is false, since there are no unicorns. Thus, if we do not assume existential commitment, an A statement will not entail an I statement with the same subject and predicate terms.

Due to similar reasoning, an E statement will entail an O statement with the same terms P and Q just in case existential commitment is assumed, and not otherwise.

This is not the only significance of existential commitment. Consider an A and an E statement with the same terms P and Q. For instance, *All giraffes are herbivores* and *No giraffes are herbivores*. Clearly, if we assume giraffes exist, then one of these statements must be false. What if we don’t make this assumption? Well, consider *All unicorns have horns* and *No unicorns have horns*. Suppose that unicorns do not exist. Then, of course, no unicorns have horns. But, just as well, all the unicorns have horns – all zero of them! So, an A and an E statement with the same terms P and Q cannot both be true if we assume existential commitment, but they can both be true if we do not. When existential commitment is assumed, we say that A and E statements are **contraries**.

Due to similar reasoning, if existential commitment is assumed, I and O statements with the same terms P and Q cannot both false. They can both be false, however, if existential commitment is not assumed. When existential commitment is assumed, we say that I and O statements are **subcontraries**.

All these relations are represented by the **classical square of opposition**:



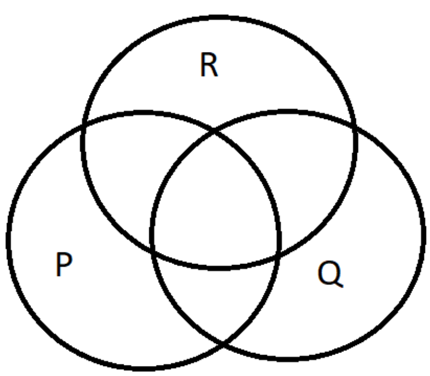
Comparing the modern and the classical squares of opposition, you can see that existential commitment makes a lot of difference. But which square is correct? The answer is: it depends on whether our statements are existentially committing! Logic tells us which statements follow from which, what combinations of statements can all be true, and what combinations of statements cannot all be true. Except in special cases, logic does not tell us when a statement is true, and logic never tells us what a statement means. Whether a statement is existentially committing or not is outside of what logic can tell us. We must look elsewhere to determine if a group of statements is existentially committing or not, and then apply logic afterwards.

Work Problems 4.1

1. Explain why, if we assume existential commitment, an E statement will entail an O statement with the same terms P and Q.
2. Explain why, if we do not assume existential commitment, an E statement will not entail an O statement with the same terms P and Q.
3. Explain why, if we assume existential commitment, a pair of I and O statements with the same terms P and Q cannot both be false.
4. Explain why, if we do not assume existential commitment, a pair of I and O statements with the same terms P and Q can both be false.

Arguments with Three Terms

So far, we have only looked at arguments in categorical logic with two terms, P and Q. When testing these arguments for validity, we used Venn diagrams with two circles: one for each term. In this section we will consider arguments containing three terms: P, Q, and R. To test three-term arguments for validity, we use Venn diagrams with three circles:

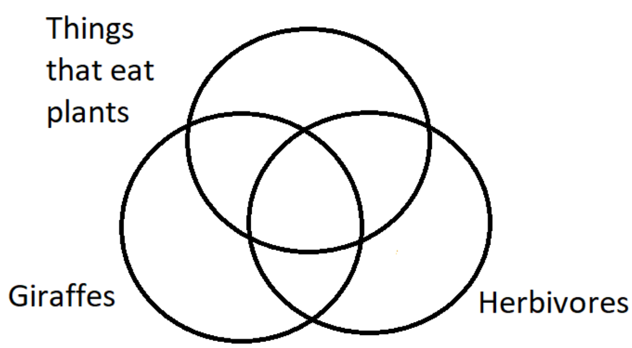


As before, each circle represents everything which falls under that category. The middle section, where all three circles overlap, represents everything which is a P, a Q, and an R. The section on the left side, where just the P and R circles overlap, represents everything which is both a P and an R. The section on the right side, where just the R and Q circles overlap, represents everything which is both an R and a Q.

Consider the following argument:

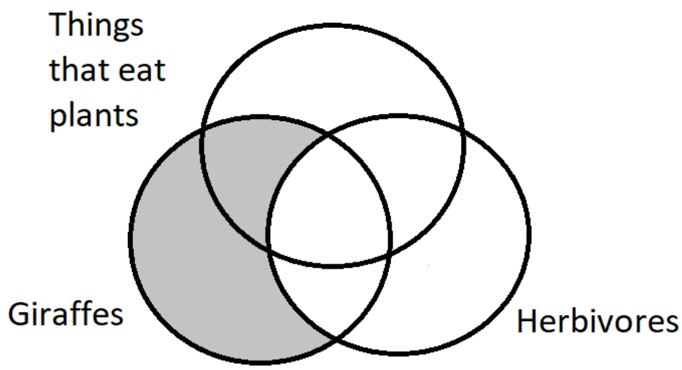
All giraffes are herbivores.	(Premise)
All herbivores are things that eat plants.	(Premise)
All giraffes are things that eat plants.	(Conclusion)

To test this argument for validity, we begin by taking our three-term Venn diagram, replacing P with "giraffes," Q with "herbivores," and R with "things that eat plants." So, we have:

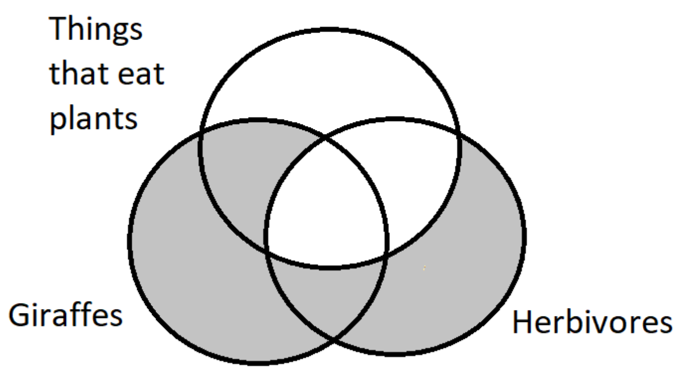


Next, we represent each premise as being true.

We can begin with the first premise, "All giraffes are herbivores." We represent this statement as being true by shading in every part of the *Giraffes* circle except the part which overlaps with the *Herbivores* circle. Remember, we use shading to show that a region is empty: there are no giraffes which fail to be herbivores. Thus, we have:

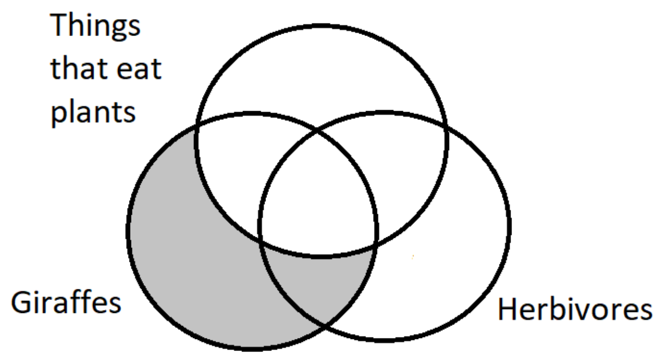


Now we represent the second premise, “All herbivores are things that eat plants,” by shading in every part of the *Herbivores* circle except what overlaps the *Things that eat plants* circle. This gives us:



This Venn diagram represents both premises being true.

Finally, we check to see if the diagram already represents the conclusion as true. The conclusion, recall, is “All giraffes are things that eat plants.” This is represented as true on the diagram by shading in every part of the *Giraffes* circle except the part which overlaps the *Things that eat plants* circle. That is:



Notice that every region which is shaded on the Venn diagram representing the conclusion is also shaded in the Venn diagram representing the two premises. Thus, by representing the two premises as being true, we automatically represented the conclusion as being true. So, the argument is valid.

You will notice that some regions which are shaded in the diagram representing the premises but are not shaded in the diagram representing the conclusion. That is fine, it just means the premises together contain more information that the conclusion. In order for an argument to be valid, the truth of the premises must guarantee the truth of the conclusion. It is fine if the truth of the premises also guarantees that some statement other than the conclusion is true as well.

We have found that the following argument is valid.

All giraffes are herbivores. (Premise)

All herbivores are things that eat plants. (Premise)

All giraffes are things that eat plants. (Conclusion)

In showing that this argument was valid, we did not consider the meaning of any of our terms. Thus, we see that

All P are Q

All Q are R

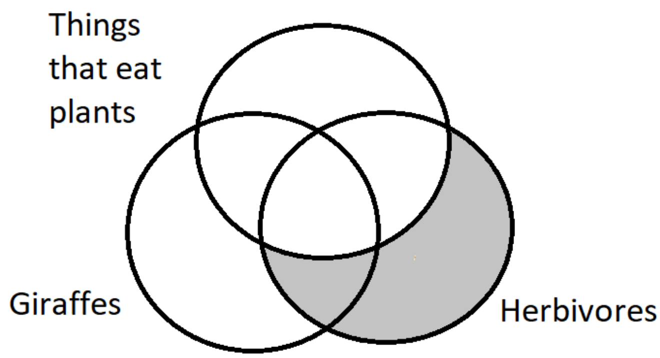
All P are R

is a valid argument form.

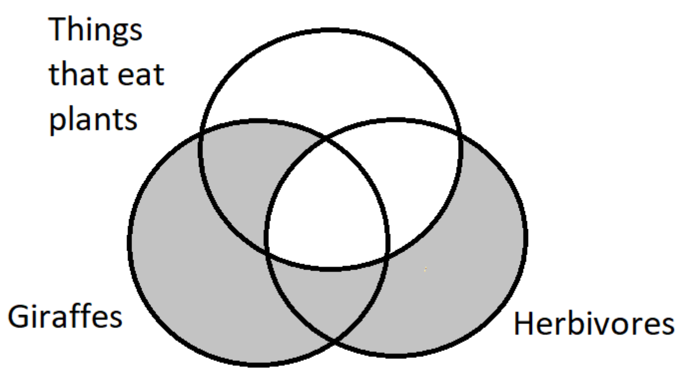
What happens if we switch the order of the two premises? Our new argument is

All herbivores are things that eat plants.	(Premise)
All giraffes are herbivores.	(Premise)
All giraffes are things that eat plants.	(Conclusion)

The first premise we represent like so:



Representing the second premise gives us:



This is exactly the same Venn diagram as with the first argument. Since the conclusion is the same in both arguments, this argument is valid. This is an important point: the order in which the premises appear in an argument has no bearing on whether the argument is valid.

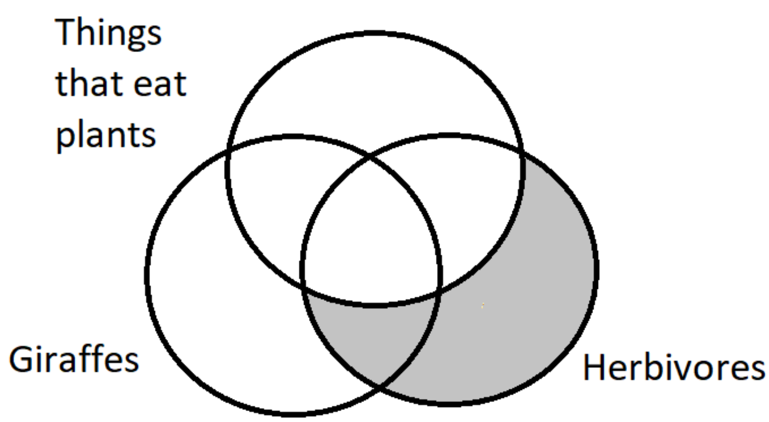
What if we switch the conclusion with one of the premises? That is, what if we make the conclusion a premise, and make one of the premises the conclusion? For instance:

All herbivores are things that eat plants. (Premise)

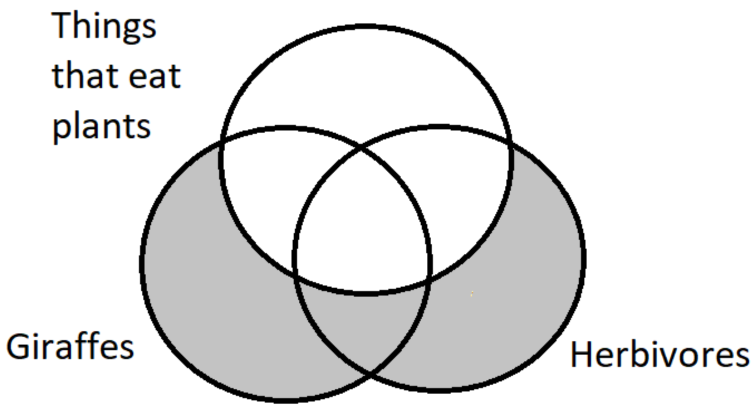
All giraffes are things that eat plants. (Premise)

All giraffes are herbivores (Conclusion)

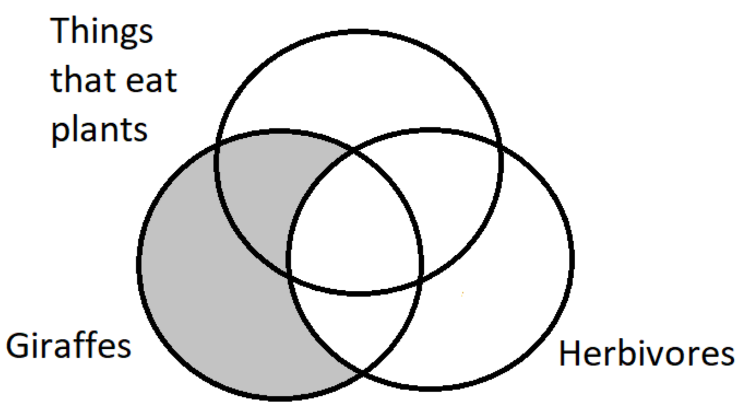
The Venn diagram representing our first premise as true looks like this:



Adding the second premise gives us this:



Now we see if this Venn diagram also represents the conclusion being true. The conclusion is "All giraffes are herbivores." Here is the Venn diagram representing the conclusion as being true:



Notice that there is a section which is shaded in the diagram representing the conclusion, but not shaded in the diagram representing the premises. Specifically, the region representing the giraffes that eat plants but are not herbivores. So, this argument is not valid. The truth of the premises does not guarantee that there are no giraffes which eat plants but are not herbivores. It is consistent with the premises that some giraffes are omnivores: animals which eat a mixture of plants and animals.

In determining that this argument is invalid, we did not consider the meaning of the terms. So, we may conclude that this

All Q are R.

All P are R.

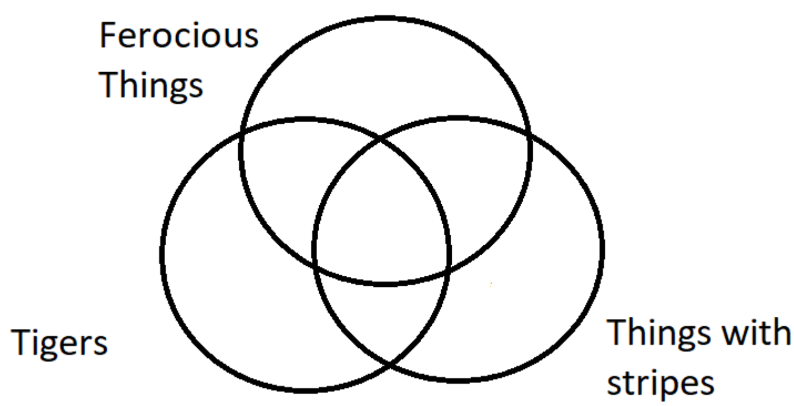
All P are Q.

is not a valid argument form.

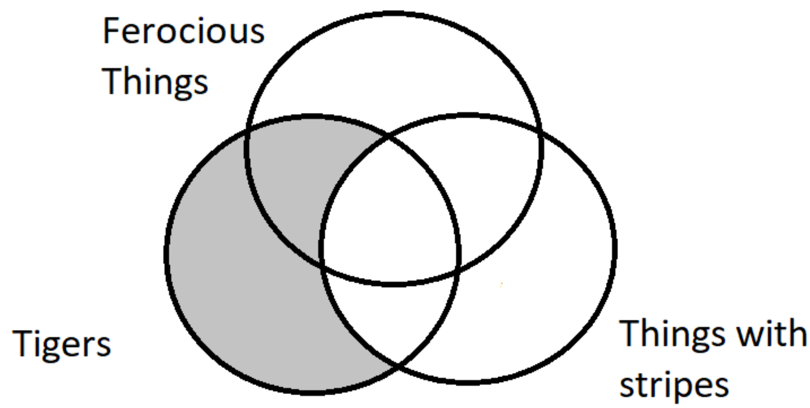
So far, our three-term arguments have used only A statements. Let us consider an argument containing an A statement and two I statements, like so:

All tigers are things with stripes.	(Premise)
Some tigers are ferocious things.	(Premise)
Some things with stripes are ferocious things.	(Conclusion)

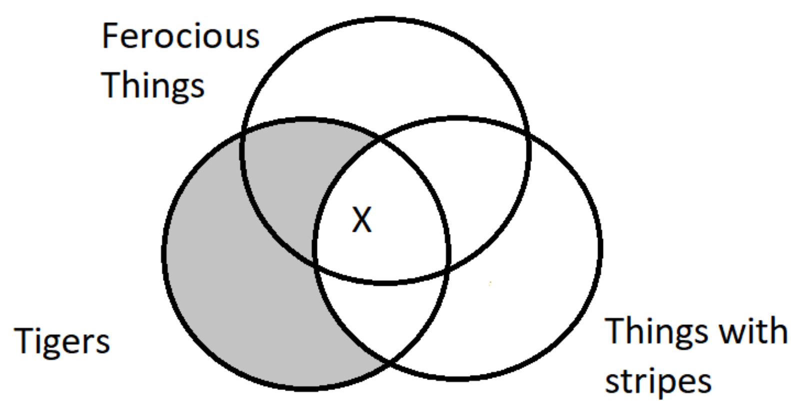
We begin with a Venn diagram like this:



We represent the first premise, "All tigers are things with stripes," like so:



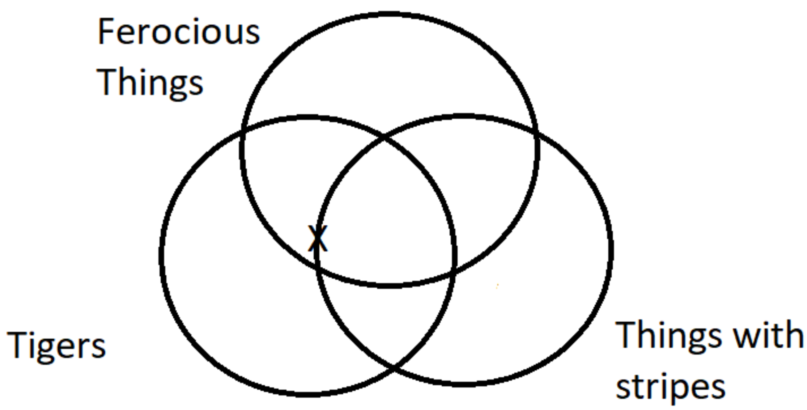
Adding the second premise, “Some tigers are ferocious things,” gives us this:



Does this Venn diagram represent the conclusion being true? Well, the conclusion, “Some things with stripes are ferocious things,” would be represented as true by a Venn diagram with an X where the circle for *Things with stripes* overlaps the circle for *Ferocious things*. Indeed, the Venn diagram representing our premises contains an X where these circles overlap. So, the argument is valid.

In representing the premises as being true, we began with the first premise. What would happen if we began with the second premise? To represent the second premise, “Some tigers are ferocious things,” as being true, we want to put an X in the section of the diagram where the

Tigers circle overlaps the *Ferocious things* circle. But there are two places where these circles overlap: the section where just those two circles overlap, and the section where every circle overlaps. Previously, when we represented the first premise first, this was not a problem, since the section where just these two circles overlap had already been shaded away. If we try to represent just the second premise, the Venn diagram looks like this:



We do not know whether to put the X in the section where just the *Tigers* circle overlaps with the *Ferocious Things* circle, or in the section where all three circles overlap. So, we put the X on the line between these sections. It is easy to see that this would mess up the diagram representing the premises. As a general rule, when making a Venn diagram to represent the premises of an argument, if possible represent A and E statements first, and I and O statements afterwards.

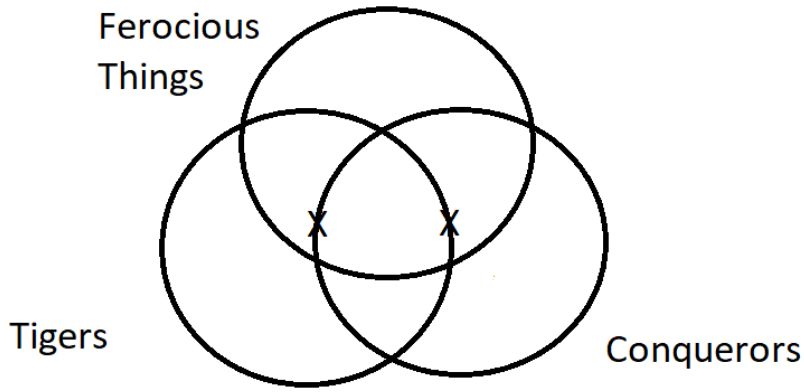
Some arguments contain no A or E statements, and so this is not possible. Consider this argument:

Some tigers are ferocious things. (Premise)

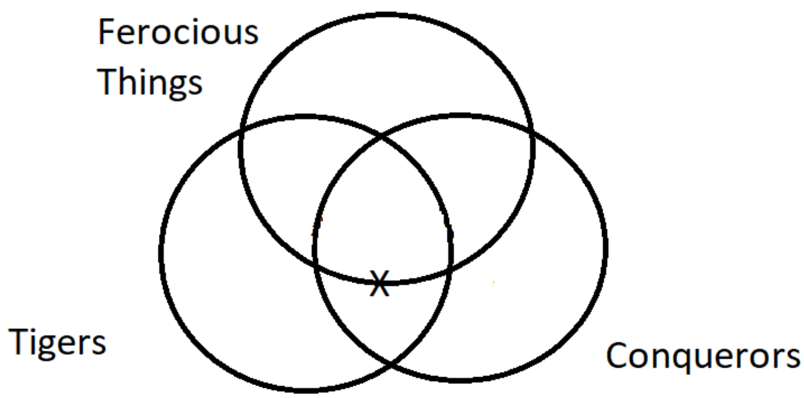
Some ferocious things are conquerors. (Premise)

Some tigers are conquerors. (Conclusion)

Here is the Venn diagram representing the premises as being true:



The Venn diagram representing the conclusion as being true looks like this:



There is clearly an X in a section of this diagram which does not appear in the diagram representing the premises. So, this argument is invalid.

Work Problems 4.2

For each of the following arguments, use a Venn diagram to show that it is valid, or give a counterexample.

1. No fish are mammals

   All furry things are mammals.

   No fish are furry things.

2. No lizards are trees.

   No trees are tigers.

   No lizards are tigers.

3. Some people are not charitable things.

   All beneficent things are charitable things.

   Some people are not beneficent things.

4. Some people are intelligent things.

   Some intelligent things are not kind things.

   Some people are not kind things.

Alterations, Compliment Classes, and Equivalencies

Each of our four basic statement types can be altered in various ways. One way is to switch the order of the subject and predicate terms. So, for instance, the A statement "All P are Q" becomes "All Q are P." Whenever you switch the order of the subject and predicate terms of a statement, the resulting statement is called the **converse** of the original:

	Original	Converse
A.	All P are Q	All Q are P

E.	No P are Q	No Q are P
I.	Some P are Q	Some Q are P
O.	Some P are not Q	Some Q are not P

An E statement and its converse, as well as an I statement and its converse, are said to be **logically equivalent**. This means that they are true in all the same possible cases. Whenever “No P are Q” is true, “No Q are P” is also true, and the other way around. Whenever “Some P are Q” is true, so is “Some Q are P,” and the other way around. If two statements S_1 and S_2 are logically equivalent, then both of the following arguments will be valid:

S_1 (Premise)

S_2 (Conclusion)

And

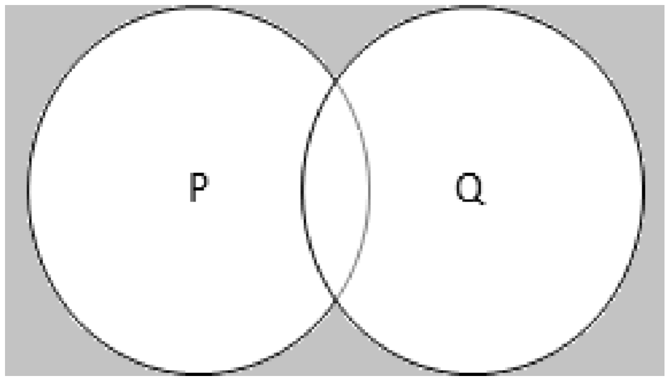
S_2 (Premise)

S_1 (Conclusion)

This is because if either statement is true, the other one must be true as well. This means that we can check for logical equivalence using Venn diagrams. An A statement and its converse are not logically equivalent, nor are an O statement and its converse.

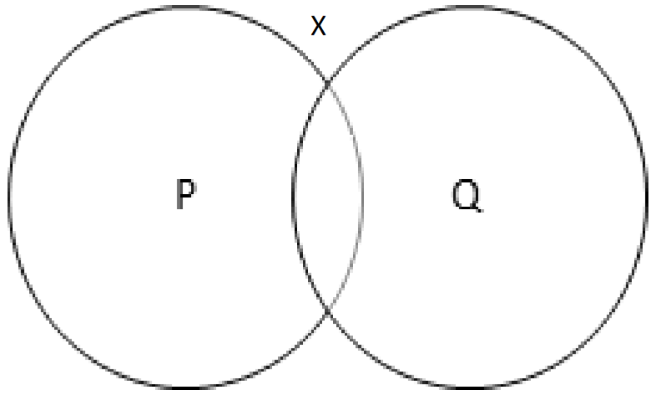
Other ways of altering our four basic statement types require introducing the notion of a **compliment class**. A **class** is just a collection of things. The compliment class of a class is simply the class of all things which are not members of the original class. So, the compliment class of P, which we may call *non-P*, is the class of everything which is not a P. One way we can modify any of our four basic statement types, is to replace either the subject or predicate term, or both, with the term for its compliment class.

For instance, suppose we replace the subject class of the A statement, “All P are Q,” with its compliment class. This gives us: “All non-P are Q.” To represent this statement on a Venn diagram, we want to shade everything which is not in the P section except what is in the Q section. Thus, we have:



This might look strange, since we are used to shading and placing Xs within the circles of the diagram. But think about what the area outside the circles on this Venn diagram represents: Everything which neither falls under the category P or the category Q. If the statement “All non-P are Q” is true, then nothing which fails to fall under the category P will also fail to fall under the category Q; that is, everything which is not a P is a Q. So, the region for anything which is neither a P nor a Q will be empty.

Consider also what happens if we take an I statement, “Some P are Q” and replace both the subject and the predicate terms with their respective compliments. This gives us, “Some non-P are non-Q.” Here is the Venn diagram:



“Some non-P are non-Q” is true just in case there is something which does not fall under either the P or Q categories. As such, our diagram has an X outside of both the P and Q circles.

We get the **obverse** of a statement by first replacing the predicate term with its compliment, and then changing the quality of the statement, either from affirmative to negative or from negative to affirmative. Here are the results:

	Original	Obverse
A.	All P are Q	No P are non-Q
E.	No P are Q	All P are non-Q
I.	Some P are Q	Some P are not non-Q
O.	Some P are not Q	Some P are non-Q

Each statement type and its observe are logically equivalent.

We get the **contraposition** of a statement by first getting the **converse** (that is, by switching the subject and predicate terms), and then replacing each term with its compliment. Here are the results:

	Original	Contraposition
A.	All P are Q	All non-Q are non-P
E.	No P are Q	No non-Q are non-P
I.	Some P are Q	Some non-P are non-Q

O. Some P are not Q — Some non-Q are not non-P

An A statement is logically equivalent to its contraposition, and an O statement is logically equivalent to its contraposition. An E statement and its contraposition are not logically equivalent, nor is an I statement and its contraposition.

Work Problems 4.3

Use Venn diagrams to show that each of the following pairs of statements is logically equivalent.

1. An E statement and its converse.
2. An I statement and its converse.
3. An A statement and its obverse.
4. An E statement and its obverse.
5. An I statements and its obverse.
6. An O statement and its obverse.
7. An A statement and its contraposition.
8. An O statements and its contraposition.

Chapter 5. Introducing Propositional Logic

As we saw in the previous two chapters, we can use categorical logic to tell us if certain arguments are logically valid. However, not every statement is one of the four types dealt with by categorical logic, nor the variations of those types. This means that categorical logic can only be used to evaluate a significantly restricted range of deductive arguments. In this chapter we turn to a different logic, **propositional logic**. Like categorical logic, propositional logic provides us with tools for identifying argument forms, and for showing that certain arguments are logically valid. However, propositional logic allows us to evaluate an entirely different class of deductive arguments than we could with categorical logic.

The Basic Idea Behind Propositional Logic

The statements of propositional logic are not statements of a natural language like English. This was true of categorical logic as well, though less obvious: "All P are Q" is not an English sentence, since "P" and "Q" are not words. It *looks* a lot like an English sentence, though. The statements of propositional logic do not look much like English sentences at all. For that reason, propositional logic can seem intimidating when you are first learning it. Although it looks unfamiliar, the language of propositional logic is both simple and powerfully expressive.

Before introducing the language, though, it may be helpful to set the stage by explaining the basic idea behind propositional logic. To begin with, it will be useful to draw a distinction between two kinds of statements in a natural language like English: **atomic** and **molecular**. A molecular statement is a statement which has at least one other statement as one of its parts. An atomic, by contrast, is a statement which does not have any other statement as one of its parts.

Here is a list of atomic statements:

Trees are plants

Dogs are mammals

Giraffes have long necks

Toby is a turtle

Susan is a snail

Here is a list of molecular statements:

Toby is not a turtle.

Either Toby is a turtle or Toby is not a turtle.

Trees are plants and dogs are mammals.

If dogs are mammals then giraffes have long necks.

Toby is a turtle if and only if Susan is a snail.

Either Toby is not a turtle or, Susan is a snail and, if trees are plants then dogs are mammals.

Notice that each molecular statement contains at least one of our atomic statements. To generate a molecular statement from a group of atomic statements, we either take an atomic statement and add "not" or something equivalent, or we combine two or more atomic statements with words or phrases like "and," "or," and "if… then." By repeating this process, we can generate molecular statements of endless variety.

Moreover, notice that whether or not a molecular statement is true is entirely a matter of whether the atomic statements composing it are true. Consider, for example,

Trees are plants and dogs are mammals.

This is true just case both of the atomic statements within it

Trees are plants

Dogs are mammals

are true. We say that the truth-value of a molecular statement is a function of the truth-values of its constitutive atomic statements. Our molecular statement

Trees are plants and dogs are mammals

has a truth-value of true only because each of the atomic statements which compose it has a truth-value of true. Were either to have a truth-value of false, the molecular statement would also be false.

The fact that the truth-value of a molecular statement is a function of the truth-values of its constitutive atomic statements is incredibly significant. An argument is valid, again, just in case it cannot both be the case that the premises are all true and the conclusion is false. That means that, if an argument is valid, there should be no way to assign truth-values to the atomic statements appearing in the argument so that the premises all come out true and the conclusion comes out false.

Propositional logic allows us to represent the way the truth-values of an atomic statement contribute to the truth-value of the whole statement. Once we can do that, we are in a position to determine which argument forms in propositional logic are valid.

Work Problems 5.1

For each statement, indicate whether it is atomic or molecular. If it is molecular, identify the atomic statement or statements within it.

1. All dogs are mammals.
2. Some dogs are kind and some dogs are mean.

3. Some cats are not housebroken.

4. Some mammals are cats and some mammals are dogs.

5. Some fish are pets, and some other fish are not pets.

6. If you want pet fish, you should get an aquarium.

7. If you want to do well, you should be kind but stand up for yourself.

8. You are either going to win or you are going to lose.

9. You will win only if you try your best.

10. People should be treated fairly.

Introducing the Language of Propositional Logic

To evaluate an argument using categorical logic, we first convert the premises and conclusion into categorical form. To evaluate an argument using propositional logic, we first represent the premises and the conclusion of the argument in the symbolic language of propositional logic. Representing a statement or statements in a symbolic language is called **translating** those statements, and the result is a **translation** of those statements.

In this section we introduce the symbolic language of propositional logic. This will consist of a group of symbols and rules for putting those symbols together to form statements in the language. The statements themselves will not mean much, if anything, to you yet. That is fine. Later on, we will discuss how to translate English statements into the symbolic language. For now, just focus on getting familiar with the symbols themselves and with what statements in the language look like.

Our logic language contains the following kinds of symbols.

1. Statement letters: a, b, c, d, e, …., z, a’, b’, c’, …, a``, b``, …

2. Logical operators: ~, &, v, →, ←→
3. Parentheses: (,)

That is all. Some books use different symbols. This does not really matter. What is important is not what symbols are used, but what they mean.

Next, we turn to the rules for making statements in the language.

1. Any statement letter is a statement in the language.
2. Where P and Q are statements in the language, the following are also statements:
    a. ~P
    b. (P & Q)
    c. (P v Q)
    d. (P → Q)
    e. (P ←→ Q)

That is all.

Notice, importantly, that P and Q in rule 2 stand for any statements in the language at all. Suppose we begin with the statement letters a, b, and c. Using rule 2b, we can generate a new statement,

(a & b),

by taking our statements a and b, putting them together with the logical operator & between them, and putting parentheses around them. Using rule 2d, we can generate the statement

(b →c)

by taking our statements b and c, putting them together with the logical operator → between them, and putting parentheses around them. Notice that these are both statements. So, using rule 2c, we can generate yet another statement, ((a & b) v (b → c)),

by taking our statements (a & b) and (b → c), putting them together with the logical operator v between them, and putting parentheses around them. Since this is a statement as well, we may use rule 2a to generate another statement,

~((a & b) v (b → c)),

by taking our previous statement and placing the logical operator ~ on the left side. By repeatedly applying these rules in various combinations, we can form more and more complex statements.

By convention, any parentheses that appear on the outside of a statement may be dropped. Thus, the first three molecular statements from the previous paragraph could have been written like so:

a & b

b → c

(a & b) v (b → c)

The fourth statement, however, had to be written like it was,

~((a & b) v (b → c)),

Because the logical operator ~ keeps the left-most parenthesis from being on the outside of the statement. As we will see later, parentheses play a very important role in the statements of our language, but they have no significance when on the outside. Officially,

a & b

is not actually a sentence in our language, but we can use it to represent

(a & b)

which is a sentence in the language. Since these parentheses do not actually contribute anything to the statement it is harmless to drop them. The reason official statements of the language have these extra parentheses is that it simplifies the rules for sentence formation.

To determine if a string of symbols is a statement in our language, we ask if that string of symbols could have been put together using the previously listed rules for sentence. Following these rules, any statement in our language must meet the following conditions.

First, a statement may only contain symbols of the appropriate kind. Importantly, a statement in our language will not contain any symbols like

A, B, C, D, P, Q, R, …

While we did use P and Q to state the rules of statement formation, P and Q are not actually symbols in the language. Instead, they are used to represent statements in the language. Why did we do this? Well, we want symbols which could represent any statement in the language at all. Each of the symbols a, b, c, d, and so on is a particular statement in the language, so we do not want it to also represent any statement in the language.

Second, no two statement letters may be directly next to one another. So, a string of symbols containing any of the following,

ab, cd, aa, pq, abcdefg, …,

is not a statement. This is a consequence of the rules 2b – 2d. A statement may contain two statement letters only if &, v, →, or ←→ is between them.

Third, none of the logical operators &, v, →, or ←→ may be directly next to any of the others or itself. So, a string of symbols containing any of the following,

v→,, &&, v→, ←→←→,

is not a statement. This is because, by rules 2b – 2d, the logical operators &, v, →, and ←→ only appear between statements, and only one at a time.

Fourth, the ~ operator will never appear immediately to the left of any of the other operators. Thus, the following

~&, ~→, ~v

are not statements in the language. By rule 2a, the ~a operator may only appear at the beginning of a statement, including at the beginning of a statement which appears inside another statement. The other logical operators, &, v, →, and ←→ always appears inside a statement, however, and not at the beginning. It is possible, however, for a statement to contain two or more instances of the logical operator ~ next to one another. By rule 2a, if ~P is a sentence, so is ~~P, ~~~P, and so on.

Fourth, for any group of three statement letters, two must be surrounded by parentheses. Thus, a string of symbols containing any of the following

a & b v c, a → b & c,

is not a statement. This is a consequence of rules 2b – 2d.

Fifth, any statement in the language must have the same number of left and right parentheses. Parentheses are introduced by rules 2b – 2d, and they are always introduced in pairs.

Work Problems 5.2

For each of the following strings of symbols, indicate whether it is a statement in our language.

1. aBc
2. (a & b)
3. ((a v b)

4. (a → (b v ~c))

5. a vv (b & c)

6. (c ←→ d)

7. e ~& f

8. ~~e

9. e & ~f

10. ~(a & ~(b → ~(d v e)))

Symbol Meaning and Translations

We have now introduced the symbols of the language, as well as explained how those symbols may be combined to form statements in the language. Having done that, we should now talk about what those symbols mean, and how this artificial language can be used to give translations of statements in a natural language like English.

Statement Letters

The statement letters, a, b, c, d, e…, are atomic statements. Each one can be assigned either truth-value: true or false. An atomic statement in the language has no particular meaning beyond this. In this respect, it is sort of like a statement form from categorical logic: "Some P are not Q" really has no particular meaning, but we can treat it as true or false when evaluating an argument form for validity.

When translating an atomic statement in English into our language, we simply assign some statement letter to stand for that statement. So, we might make the following assignments:

Trees are plants: t

Dogs are mammals:	d
Giraffes have long necks:	g
Toby is a turtle:	r
Susan is a snail:	s

It does not really matter what statement letter we assign to each atomic statement. However, each unique atomic statement in an argument should be assigned a unique statement letter, and all instances of a single atomic statement in an argument should be assigned the same statement letter. So, all of the above statements appeared in an argument, we would not want to assign *t* to both "Trees are plants" and "Dogs are mammals." If we did, the argument in our language would treat these as the same atomic statement, which they are not. Likewise, if "Giraffes have long necks" were to appear twice in an argument, we would not want to assign *g* to one instance and *f* to another. If we did, the argument in our language would treat these as different atomic statements, which they are not.

Work Problems 5.3

For each of the following statements, identify the atomic statements within it, and assign a letter to each atomic statement.

1. All fish live in water.
2. Not all fish live in water.
3. All cats live on land.
4. Fish live in water and cats live on land.
5. Dogs bark at strangers.
6. Either dogs bark at strangers, or fish live in water or cats live on land.

7. Frogs are mammals.
8. Frogs are not mammals.
9. If frogs are not mammals, then dogs bark at strangers.
10. Either frogs are not mammals or fish live in water.

Logical Operators

The meaning of the logical operators, ~, &, v, →, and ↔, and the parentheses is the function each plays in any statement in which it appears. In this section we explain the function of the logical operators, and in the next we explain the function of the parentheses.

The ~ operator modifies any statement which it appears to the left of, and each of the other four operators combines two smaller statements into a larger statement. Whenever a logical operator and one or two statements combine to form a larger statement, the truth-value of that larger statement is determined by the truth-values assigned to the smaller statements within it together with the logical operator. Since all statements in the language are built up out of smaller statements using the logical operators (and parentheses, the function of which will be explained shortly), the truth-value of any statement in our language ends up being determined by the truth-values assigned to the atomic statements composing it together with its logical operators (and parentheses).

To explain how the truth-value of an atomic statement or pair of statements together with a logical operator determines the truth-value of the statement they compose, we will use truth-tables. A truth-table is a chart which shows what truth-value a statement has based on the truth-values assigned to its parts. For the following truth-tables, we will let P and Q stand for statements in the language.

We form a statement containing the ~ operator by placing it to the left of a statement, as in ~P. This larger statement is composed of P and the operator ~. We begin our truth-table like so:

P	~P

Here we have P above the line on the left, and ~P above the line on the right. Below the line on the left, under P, we put all the possible truth-values for P. P could either be true or false Thus:

P	~P
T	
F	

Finally, below the line on the right side, under ~P, we put the truth-values which ~P receives depending on the truth-value of P. Here is the result:

P	~P
T	F
F	T

What this table says is that when P is true, ~P is false, and when P is false, ~P is true.

Next, we turn to the other four operators. We will make a single table on which all four are represented. We begin like this:

P	Q	(P & Q)	(P v Q)	(P → Q)	(P ←→ Q)

Next, below the line on the left side we assign all the possible combinations of truth-values to our statements P and Q. Both can be true, both can be false, or one could be true and the other false. Thus:

P	Q	(P & Q)	(P v Q)	(P → Q)	(P ↔ Q)
T	T				
T	F				
F	T				
F	F				

Finally, below the line on the right side we assign truth-values to each new statement based on the truth-values of P and Q. Here is the result:

P	Q	(P & Q)	(P v Q)	(P → Q)	(P ↔ Q)
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

(P & Q) is true just in case P and Q are both true, and false otherwise. (P v Q) is true if P is true, if Q is true, or if both are true, and false if both P and Q are false. (P → Q) is true if either P is false or if Q is true. Finally, (P ↔ Q) is true just in case P and Q both have the same truth-value.

By looking at these truth-tables, you can probably make a pretty good guess as to what English terms and phrases some of these operators are used to translate. Here is a table listing some examples:

~	&	V	→	↔
not it is not the case	And but however	Or either… or unless	if … then only if	if and only if

The following table lists several molecular statements, and possible translations for them:

Toby is not a turtle	~t
Trees are plants and dogs are mammals	p & m

If dogs are mammals then giraffes have long necks	d → g
Toby is a turtle if and only if Susan is a snail	t ←→ s

Some of the words and phrases the logical operators are used to translate will make sense, but some may be confusing. It will be useful to discuss some of them. It makes sense that & would be used to translate "and," but it might come as a surprise that & also translates words like "but." People just learning propositional logic, when tasked with translating a statement containing "but" often try to use ~. This is not correct.

Consider as an example:

Jerry thought it would rain today, but the sky remained clear.

Under what conditions would this statement be true? Clearly it would be false if Jerry did not think it was going to rain today. The statement would also be false if it did in fact rain today. Obviously, it would be false if both these claims were false. Thinking about it, it is easy to see that this statement is true only if both "Jerry thought it would rain today" and "The sky remained clear" are true. The operator with these truth-conditions is &. In English, the word "but" suggests a contrast between what is the case (or what is claimed to be the case) and what is expected, whereas the word "and" does not suggest any such contrast. The language of propositional logic is only concerned with how the truth-values of the atomic parts of a statement contribute to the truth-value of the whole statement. For this purpose, the difference between "but" and "and" in English is unimportant. Thus, from the perspective of logic, "and" and "but" are equivalent.

Next consider v. For the most part, it is not surprising that v translates "or". It may seem a bit strange that a statement containing v can be true even if both parts of the statement can be true. Take a statement like

John's pet is either a dog or a lizard.

John's pet cannot be both a dog and a lizard! But consider instead a statement like:

John is either a student or a teacher.

Could John be both a student and a teacher? Of course! And if he was, would this statement be true? Again, of course. In a statement containing or, if both parts are true, the whole statement is true. This is true even in the case of "John's pet is either a dog or a lizard." It happens in this case that the two parts cannot both be true, but this has nothing to do with the meaning of "or," but rather with facts about dogs and lizards. You might wonder, though: what if we want to say that John is either a teacher or a student, but we do not want to say he is both? Well, we just add that into the statement:

John is either a teacher or a student, but he is not both.

Later we will discuss how to translate a statement like this into our symbolic language.

The operator v is also used to translate the word "unless." The situation here is much like that with "but" being translated by &. Consider, for example, the statement

Jill will go see a movie alone unless Tom invites her to the party.

Under would conditions would this statement be false? Well, it would be false if Tom does not invite Jill to the party and Jill does not go see a movie alone. Here are the other possible cases

Jill goes to see the movie and Tom invites her to the party.

Jill does not go to see the movie and Tom invites her to the party.

Jill goes to see the movie and Tom does not invite her to the party.

The original statement will be true in each of these cases. One might wonder about the first case. Notice, however, that the statement

Jill will go see a movie alone unless Tom invites her to the party, but she might go to the movie even if he does, is perfectly consistent. So, although someone who says

Jill will go see a movie alone unless Tom invites her to the party

may well be suggesting that Sara will not go the movies if Tom invites her to the party, that is not strictly part of the literal meaning of the statement. In English, the word "unless" differs from "or" in that a statement of the form "P unless Q" is often used to suggest that "P" is a kind of default, with "Q" being an exception. Someone who says "Jill will go see a movie alone unless Tom invites her to the party" likely intends to express that it is more likely that Jill sees a movie alone than not. The word "or" does not function this way. This difference is ignored in the language of propositional logic, which focuses on the way the truth-values of atomic statements contribute to the truth-values of molecular statements which they help compose. So, propositional logic treats "or" and "unless" as equivalent.

This leaves us with → and ←→. We began with →. Consider the following statements:

If Joaquim is kind, then Monique is smart.

If Joaquim is kind, Monique is smart.

Joaquim is kind only if Monique is smart.

Monique is kind if Joaquim is smart.

Each of these statements means exactly the same thing. Each of these is an example of a **material conditional**. The two statements which compose a material conditional have special names: the **antecedent** and the **consequent**. Here the antecedent and consequent are identified for each of the above ways of writing a material conditional:

If antecedent, then consequent.

If antecedent, consequent.

Antecedent only if consequent.

Consequent, if antecedent.

In a conditional statement, we say that the antecedent is **sufficient** for the consequent, and the consequent is **necessary** for the antecedent. That the antecedent is sufficient for the consequent means that, if the antecedent is true the consequent must be true as well, assuming the conditional is true. That the consequent is necessary for 0the antecedent meant that the consequent must be true for the antecedent to be true, again assuming the conditional is true.

A material conditional is true just in case either the antecedent is false, or the consequent is true. This means that a material conditional is false just in case the antecedent is true and the consequent is false. Here is one way to see this. Consider the statement

If Ralph is a dog, then Ralph is a mammal.

This is equivalent to

Either Ralph is not a dog, or Ralph is a mammal.

The latter statement contains an "or", so it is true just in case either "Ralph is not a dog" is true or "Ralph is a mammal" is true or both. Since the statement "Ralph is not a dog" contains a "not," it is true just in case "Ralph is a dog" is false. So,

Either Ralph is not a dog, or Ralph is a mammal

is true just in case either "Ralph is a dog" is false or "Ralph is a mammal" is true. Then, since the statements are equivalent

If Ralph is a dog, then Ralph is a mammal

is true just in case either "Ralph is a dog" is false" or "Ralph is a mammal" is true. That is, it is true just in case either the antecedent is false, or the consequent is true.

Part of the reason the truth-conditions for material conditionals can seem surprising is that we might confuse them with other kinds of conditional statements. Take the statement

If I were to drop the ball, it would fall to the floor.

The antecedent of this statement is false just in case I do not drop the ball. Suppose I do not drop the ball. Is this enough to make the statement true? Now, obviously the statement is true, but that is not what we are asking. Rather, we are asking if the reason the statement is true is that I do not drop the ball. The answer is no. To see this, consider a different statement,

If I were to drop the ball, it would float for five seconds and then shoot up to the ceiling.

This statement is clearly false and will be false regardless of whether I drop the ball.

Statements like "If I were to drop the ball, it would fall to the floor" and "If I were to drop the ball, it would float for five seconds and then shoot up to the ceiling" are called **counterfactual conditionals**. Material conditionals and counterfactual conditions are different, as we can see. Propositional logic lacks the tools to translate counterfactual conditionals. To translate counterfactual conditionals, we need **modal logic**, which deals with statements about possibility and necessity. Modal logic is incredibly fascinating, but it goes beyond what we will cover in this book.

This leaves us with ←→, which is used to translate the phrase "if and only if." Such statements are called **biconditionals**. A biconditional is really two conditional statements hooked together. For example, the statement

Joaquim is kind if and only if Monique is clever

is equivalent to

Joaquim is kind if Monique is clever, and Monique is clever if Joaquim is kind,

which is the same as

If Monique is clever then Jamal is kind, and if Jamal is kind them Monique is clever.

Here we have two statements joined by “and,” which means the whole statement is true if both parts are true, and otherwise false. If “Monique is clever” is true but “Jamal is kind” is false, then “If Monique is clever than Jamal is kind” is false, and so the biconditional is false. If “Jamal is kind” is true but “Monique is clever” is false, then “If Jamal is kind then Monique is clever” is false, and so the biconditional is false. If both “Monique is clever” and “Jamal is kind” are true, both conditionals are true because the consequent of both is true, and so the biconditional is true. If both “Monique is clever” and “Jamal is kind” are false, then both conditionals are true because the antecedent of both is false. So, a biconditional is true if both statements which compose it are true, or if both statements which compose it are false.

Work Problems 5.4

1. Translate the following statements into our logical language.
    1.1. Monique is great a math.
    1.2. Monique is not great at math.
    1.3. Charlie is kind.
    1.4. Charlie is not kind.
    1.5. Either Monique is great at math or Charlie is kind.
    1.6. If Charlie is kind, then Monique is great at math.
    1.7. Charlie is kind if and only if Monique is not great at math.
    1.8. If Charlie is not kind, then Monique is not great at math.
    1.9. Either Monique is not kind or Charlie is great at math.

2. Let “xor” be just like “or”, except a statement containing “xor” is false if both parts are true. Make a truth-table for “xor”.

Parentheses

Our language has only one more kind of symbol: parentheses. To explain the function of parentheses, it will be useful to introduce them through an example. Compare the following two statements. First,

Either Toby is a turtle and Gerry is a giraffe, or Ralph is a dog.

Second,

Toby is a turtle and either Gerry is a giraffe or Ralph is a dog.

How might we symbolize these statements? First, we could assign statement letters to the atomic parts like so:

Toby is a turtle:	t
Gerry is a giraffe:	g
Ralph is a dog:	d

Since these are molecular statements, we need to use our logical operators. The first statement contains an “and” and an “or,” so we will want to use both & and v. We might try:

t & g v d.

This is wrong. Remember, any statement with three statement letters must have two of them in parentheses. But we will ignore that for the moment. What about the second statement. It has all the same atomic statements, and also contains “and” and “or,” and in the very same order. Thus, again we have

t & g v d.

Again, we ignore the fact that this is not really a statement in the language.

Our attempted representations of these statements are exactly the same. But are these statements logically equivalent? That is, are they true in all the same cases? They are not. For one thing, the first statement would be true if Ralph was a dog, but this by itself would not make the second statement true.

Parentheses allow us to represent these different statements differently. The first statement

Either Toby is a turtle and Gerry is a giraffe, or Ralph is a dog,

can be translated like so:

(t & g) v r.

The second statement

Toby is a turtle and either Gerry is a giraffe or Ralph is a dog

can be translated like this:

t & (t v r).

Parentheses place different parts of a statement into groups, and the way the parts of a statement are grouped can impact the meaning of the whole statement.

As another example of where parentheses are necessary for translation, consider the following pair of statements. First,

If it is not the case that Toby is a turtle, then Ralph is a dog.

This statement is translated as follows:

~t → r.

Second,

It is not the case that, if Toby is a turtle, then Ralph is a dog.

This statement is translated as:

~(t → r).

As in our previous example, the translations of these statements contain the very same statement letters and the very same logical operators, all appearing in the same order. Yet the two English statements are clearly not equivalent. The difference is captured by the placement of the parentheses.

We have seen that the placement of parentheses within a statement effects the meaning of that statement. But how does it do this? To assign truth-values to a molecular statement, you begin by assigning truth-values to the atomic statements, and then you assign truth-values to the molecular statements within the statement, until eventually you are able to assign a truth-value to the statement as a whole. A pair of parentheses functions to organize the parts of a molecular statement into groups, where whatever falls within a pair of parentheses is a statement. Thus, two molecular statements with the same tense operators, in the same order, but with different parentheses will be composed of different statements, and hence may have different meanings. So, in

(t & q) v r,

"t & q" is a statement within the whole statement, whereas in

t & (q v r),

"q v r" is a statement within the whole statement. In the first case, we assign a truth-value to "t & q" before assigning a truth-value to the whole statement, and the in second case we first assign a truth-value to "q v r."

We can see this with truth-tables. Let us start with "~t → r" and "~(t → r)." We begin like this:

t r	~t → r	~(t → r)

Our statement letters, t and r, are on the top left of the table, and our two molecular statements are on the top right. Next, we want to assign all possible combinations of truth-values to our statement letters: both true, both false, and one true and the other false. This gives us:

t r	~t → r	~(t → r)
T T		
T F		
F T		
F F		

Next, we need to assign truth-values to our operators. But in what order? We can start with "~t → r." The ~ operator modifies whatever statement appears on its left. In this case, t. The → operator combines the statements on either side of it. In this case, these are ~t and r. This means that we cannot determine what truth-values to assign to → until we know what truth-values to assign to ~t. So, we start there.

We know from earlier that ~P always has the opposite value as P. So, in this case ~t will have the opposite truth-values as t. Thus:

t r	~t → r	~(t → r)
T T	F	
T F	F	
F T	T	
F F	T	

Finally, we can assign truth-values to →, which is true just in case the antecedent is false, or the consequent is true. So:

t r	~t → r	~(t → r)
T T	F T	
T F	F T	
F T	T T	
F F	T F	

Now we have the truth-table for "~t → r." It is true in every case except when t and r are both false. The column showing the truth-values for the whole statement has been circled for clarity.

Let us turn to "~(t → r)." The ~ operator modifies whatever statement it appears on its left. In this case, this is "(t → r)." The → operator combines the statements on either side of it, in this case, t and r. We cannot assign truth-values to ~ until we have assigned truth-values to →. So, we start there, remembering that a conditional is true just in case the antecedent is false, or the consequent is true. Hence:

t r	~t → r	~(t → r)
T T	F T	T
T F	F T	F
F T	T T	T
F F	T F	T

Finally, we can assign truth-values to ~, recalling that ~P always has the opposite truth-value as P. So:

t r	~t → r	~(t → r)
T T	F T	F T
T F	F T	T F
F T	T T	F T
F F	T F	F T

Now we have the truth-table for "~(t → r)." It is false in every case except when t is true and r is false. The column showing truth-values for the whole statement has been circled for clarity. We see that "~t → r" and "~(t → r)" are true in different cases.

Next, we can compare "(t & q) v r" and "t & (q v r)." We start with our statement letters on the top left of our table, and the statements on the top right:

t q r	(t & q) v r	t & (q v r)

Next, we assign all possible combinations of truth-values to our statement letters. All three can be true, all three can be false, any two can be true, and any two can be false. This gives us:

t q r	(t & q) v r	t & (q v r)
T T T		
T T F		
T F T		
T F F		
F T T		
F T F		
F F T		
F F F		

Now we are ready to start assigning truth-values to our statements.

(By the way, there is a trick for assigning truth-values to atomic statements. We know that each atomic statement can either be true or false. On each row of the truth-table, each atomic statement will take one these truth-values. Where n = the number of statement letters, the number of columns = 2^n. So, with 1 statement letter (as in ~P), we have $2^1 = 2$ rows. With 2 statement letters, we have $2^2 = 4$ rows. With 3 statement letters we have $2^3 = 2 \times 2 \times 2 = 8$ rows. Take the number of rows and divide it in half. So, 4 if you have 8 rows. In the first column, make that many starting with the top true, and the rest false. Thus, 4 true then 4 false for the first column, assuming 8 rows. Next take the number of rows and divide it in half twice. In the second column, starting from the top, alternate making that many true, then false. So, if you have 8 rows, two true, two false, two true, and two false. Next take the number of rows and divide it in half again. In the third column, alternate that many true and false. Continue until you have filled every column. This will give you all the possible combinations.)

Let's begin with "(t & q) v r." Both & and v combine the two statements around them. In this case, v combines "(t & q)" and "r." So, we cannot assign truth-values to v until after we have assigned truth-values to &. "P & Q" is true just in case P and true and Q is true, and false otherwise. So

t q r	(t & q) v r	t & (q v r)
T T T	T	
T T F	T	
T F T	F	
T F F	F	
F T T	F	
F T F	F	
F F T	F	
F F F	F	

Finally, we assign truth-values to v. "P v Q" is true just in case P is true, or Q is true, or both. So,

t q r	(t & q) v r		t & (q v r)
T T T	T	T	
T T F	T	T	
T F T	F	T	
T F F	F	F	
F T T	F	T	
F T F	F	F	
F F T	F	T	
F F F	F	T	

The column with truth-values for the whole statement has been circled for clarity.

Turning to "t & (q v r)," we see that we cannot assign truth-values to & until after assigning truth-values to v. So:

t q r	(t & q) v r		t & (q v r)
T T T	T	T	T
T T F	T	T	T
T F T	F	T	T
T F F	F	F	F
F T T	F	T	T
F T F	F	F	T
F F T	F	T	T
F F F	F	T	F

Finally, we assign truth-values to &, as such:

t q r	(t & q) v r		t & (q v r)	
T T T	T	T	T	T
T T F	T	T	T	T
T F T	F	T	T	T
T F F	F	F	F	F
F T T	F	T	F	T
F T F	F	F	F	T
F F T	F	T	F	T
F F F	F	T	F	F

Again, the column showing the truth-values for the whole statement has been circled for clarity. Looking at our truth-table, we can clearly see that "(t & q) v r" and "t & (q v r)" are true in different cases. That is, they have different **truth-conditions**.

Whenever translating statements from a natural language, like English, into our symbolic language, it is essential to think carefully about the structure of the statement. How is this statement built up out of smaller statements? Consider,

If trees are plants and dogs are mammals, then, giraffes have long necks and either Toby is a turtle or Susan is a snail.

This statement contains five atomic parts. How are they put together? Well, notice first that in "If trees are plants and dogs are mammals, then" we have an "and" between "if" and "then". This means that

Trees are plants and dogs are mammals

is the antecedent of a material conditional. What is the consequent? Looking at what occurs after "then," we notice that

Toby is a turtle or Susan is a snail

is grouped together by the word "either." So, the consequent of our conditional is either

Giraffes have long necks,

or it is

Giraffes have long necks and either Toby is a turtle or Susan is a snail.

Well, which is it? Think about how much of what comes after "then" is a part of the clause introduced by "then." Doing this, we can see that the consequent is

Giraffes have long necks and either Toby is a turtle or Susan is a snail.

So, we translate our statement like so:

(t & d) → (g & (t v s)).

The parentheses and operators together nicely show the way the atomic statements come together to compose the whole statement.

We decided that, in the previous statement, everything after “then” is part of the consequent. But what if we wanted the consequent to just be

Giraffes have long necks,

and for

Either Toby is a turtle or Susan is a snail

to be a different part of the statement? In that case, we might write the statement like this:

If trees are plants and dogs are mammals then giraffes have long necks, and either Toby is a turtle or Susan is a snail.

Here is how this statement should be symbolized:

((t & d) → g) & (t v s)

For comparison, here is the original, and its symbolization

If trees are plants and dogs are mammals, then, giraffes have long necks and either Toby is a turtle or Susan is a snail.

(t & d) → (g & (t v s))

In many cases, like this one, it is crucial to pay attention to punctuation when doing translations.

Before closing this section, we will make good on our earlier promise to discuss the statement

John is either a teacher or a student, but he is not both.

How do we go about translating this statement? To begin, we identify the atomic statements within it, and assign them statement letters. Thus:

John is a teacher t

John is a student s.

Notice that neither of these atomic statements actually appears in the original. Natural languages like English often have ways of abbreviating statements. When doing translations, it is often helpful to rewrite the statement so as to make every instance of each atomic statement explicit. Here is the result of doing so:

John is a teacher or John is a student, but it is not the case that John is a teacher and John is a student.

Now we ask how the atomic statements within the statement are grouped together. The comma after "student" suggests that

John is a teacher or John is a student

Forms a group. The phrase "it is not the case" would be symbolized by ~, which always modifies a single statement. In this case, ~ modifies

John is a teacher and John is a student.

So, it looks like this statement has two main parts

John is a teacher or John is a student

and

It is not the case that John is a teacher and John is a student,

which are combined with the word "but." Recalling that "or" is symbolized with v, and "and" and "but" with &, here is the translation:

(t v s) & ~(t & s).

Here is the truth-table:

t s	(t v s) & ~(t & s)
T T	T F F T
T F	T T T F
F T	T T T F
F F	F F T F

The column showing the truth-values for the whole statement has been circled for clarity. As we can see, the original statement is true if John is a teacher, it is true if John is a student, but it is not true if John is both a teacher and a student. This is just what we should expect.

Work Problems 5.5

For each of the following statements, give a translation in our symbolic language, and make a truth-table.

1. Some fish are pets, and some fish are not pets.
2. If you want to do well, you should be kind but stand up for yourself.
3. Either dogs bark at strangers, or fish live in water or cats live on land.
4. Either frogs are not mammals or fish live in water.
5. If you learn about deductive and inductive reasoning, you will become a better critical thinker.

Chapter 6. Statement Types and Evaluating Arguments in Propositional Logic

In this chapter, we continue our discussion of the language of propositional logic by discussing different ways of classifying statements in propositional logic. Afterwards, we turn our attention to arguments in propositional logic, discussing logical form and two methods for assessing arguments in propositional logic for logical validity.

Molecular Statement Types

You will recall that, in categorical logic, we had our four basic statement forms, plus their various alterations. We can also classify statements in propositional logic, and in fact there are multiple ways of doing so, which we will discuss in the three sections to follow.

Statement Types and Main Operators

One way of classifying molecular statements is by **statement type**. To find the statement type of a molecular statement, you must identify the **main operator** of the statement. As we have seen, whenever we assign truth-values to a statement on a truth-table, we must assign those truth-values in the correct order. For any molecular statement, it will always be the case that some single operator is assigned truth-values last: this is the main operator.

The following table gives the statement type associated with each main operator:

Main Operator	Statement Type
~	Negation
&	Conjunction
V	Disjunction
→	Conditional
←→	Biconditional

While statements in our symbolic language can be enormously complex and varied, every molecular statement will be an instance of one of just these five **statement types**.

The following statement

It is not the case that, if Toby is a turtle, then Ralph is a dog.

is a negation. These statements

John is either a teacher or a student, but he is not both.

Trees are plants and dogs are mammals.

Joaquim is kind if Monique is clever, and Monique is clever if Joaquim is kind.

Toby is a turtle and either Gerry is a giraffe or Ralph is a dog.

If trees are plants and dogs are mammals then giraffes have long necks, and either Toby is a turtle or Susan is a snail.

are all conjunctions. These statements

Either Toby is not a turtle or, Susan is a snail and, if trees are plants then dogs are mammals.

Either Toby is a turtle and Gerry is a giraffe, or Ralph is a dog.

are both disjunctions. These statements

If dogs are mammals then giraffes have long necks.

If it is not the case that Toby is a turtle, then Ralph is a dog.

If trees are plants and dogs are mammals, then, giraffes have long necks and either Toby is a turtle or Susan is a snail.

are all conditionals. Finally,

Toby is a turtle if and only if Susan is a snail.

Joaquim is kind if and only if Monique is clever.

are both biconditionals.

The nice thing about classifying statements in this way is that any given molecular statement will have only one main operator, so that each molecular statement will have exactly one statement type. Furthermore, we know that every negation, no matter how complicated, is true just in case the statement being negated is false, and false just in case the statement being negated is true. We know that every conjunction, no matter how complicated, is true just in case the statements on both sides of the main operator are true, and it is false otherwise. And so on for the remaining operators.

Work Problems 6.1

For each of the following statements, identify the statement type.

1. Jane is either smart or kind.
2. If Justin is working on his novel, Rachel is working on her screenplay.
3. It is not the case that Tom is doing a trick and Sam is telling a joke.
4. Tom is not doing a trick, and Sam is telling a joke.
5. It is either the case that you will if you do your best, or that you won't win.
6. Vanessa likes horses and either Jim is an astronaut or a philosopher.
7. If fish swim, they either live in fresh water or salt water.

8. If Phil's pet is either a cat or a dog, it is a mammal.

9. Some birds fly, but some do not.

10. It is not the case that all birds fly or that all pets are mammals.

Statement Forms and Substitution Instances

Another way to classify statements is by **statement form**. To find a form of a statement, simply translate that statement into our logical language. Thus,

Either Jane is kind or Jamal is smart,

has

k v s.

as a statement form.

We say *a* statement form, because in fact every molecular statement will have multiple statement forms. To see that this is the case, we must first talk about **substitution instances**. A substitution instance is the result of taking a statement, say

p → q

and replacing every instance of some statement letter or letters with any statement in the language. For instance, if we replace p with

t & d

and we replace q with

g & (t v s),

the result is

(t & d) → (g & (t v s))

which is a substitution instance of p → q.

Since we can replace p and q with anything we like, the number of substitution instances of a given statement are potentially infinite. For instance, the following statements

$\sim((a \vee b) \,\&\, c) \rightarrow (d \leftrightarrow \sim(e \vee (f \rightarrow g)))$

$((p \rightarrow (q \,\&\, r)) \rightarrow \sim(s \vee (t \,\&\, (u \rightarrow \sim v)))$

are both substitution instances of $p \rightarrow q$.

Just as we can take a statement form and generate a substitution instance by replacing the statement letters with statements, we can take a statement and find a statement form of which it is a substitution instance by replacing the statements within it with statement letters. So, take the following statement:

$((p \rightarrow (q \,\&\, r)) \rightarrow \sim(s \vee (t \,\&\, (u \rightarrow \sim v)))$.

Next, let us make the following replacements

$(q \,\&\, r)$	a
$(u \rightarrow \sim v)$	b.

This gives us:

$(p \rightarrow q) \rightarrow \sim(s \vee (t \,\&\, b))$.

Next, let us make the following, further replacements:

$(p \rightarrow q)$	c
$(t \,\&\, b)$	d.

This gives us:

$c \rightarrow \sim(s \vee d)$.

Next, we make the following replacement:

$(s \vee d)$	e.

This gives us:

c → ~e.

We may also make this replacement:

~e   f.

This gives us

c → f.

Finally, we may make yet another replacement:

c → f   g.

This gives us

g.

Our original statement,

((p → (q & r)) → ~(s v (t & (u → ~v)),

is a substitution instance of each of the following statements:

(p → q) → ~(s v (t & b))

c → ~(s v d)

c → ~e

c → f

g

Every substitution instance of a statement is a statement form of that statement. Any molecular statement will be a substitution instance of at least two statements (one will be just a statement letter, and one will contain exactly one logical operator). So, any molecular statement will have at least two statement forms.

When finding a statement form of a statement in our logical language, you must replace statements with letters. Consider again the statement ((p → (q & r)) → ~(s v (t & (u → ~v)).

Notice that

(q & r) → ~(s

is not a statement. This means that we cannot replace it with a letter. So,

((p → (q & r)) → ~(s v (t & (u → ~v))

is not a substitution instance of

(a v (t & (u → ~v)).

Suppose we take a molecular statement and start replacing its parts with letters, and we continue until we have a statement form with just a single logical operator. It turns out, this will be the main operator of the statement.

Why is this the case? Well, we cannot replace a statement with a letter unless we replace the whole statement. If one statement appears within another statement, the smaller statement is a part of the larger statement. If the main operator of a statement is ~, any other operators will be within the statement which ~ modifies. If the main operator of a statement is any of &, v, →, or ←→, any other operators will be within one of the two statements which the main operator combines. Thus, any operator other than the main operator will be a part of a statement which the main operator either modifies or combines with another statement. Hence, if we replace parts of a statement with letters, so that the main operator is replaced but some other operator is not, we will have to replace a statement with a letter without replacing a part of that statement. But that would mean replacing merely part of a statement with a letter, which we cannot do. So, we cannot replace a statement with the main operator with a letter and leave some other operator to remain. Hence, if we replace all the operators but one, that operator must be the main operator.

The reasoning in the previous paragraph may be a bit hard to follow. Here is the short version. When substituting letters for statements, the only way to replace a statement containing the main operator and leave another operator unreplaced is to replace only part of a statement. For instance, take the statement

$$((p \rightarrow (q \& r)) \rightarrow {\sim}(s \vee (t \& (u \rightarrow {\sim}v)).$$

I get this statement

$$a \rightarrow {\sim}v))$$

by replacing

$$((p \rightarrow (q – r)) \rightarrow {\sim}(s \vee (t \& (u$$

with “a.” But, the latter is not a whole statement, and when you replace statements with letters you must replace whole statements.

In any case, if we replace statements with letters in a molecular statement so that only one operator remains, that will be the main operator. So, one way of uniquely identifying the form of an atomic statement, is by symbolizing the statement type.

Work Problems 6.2

For each of the following statements, identify the statement form corresponding to the statement type, and give one substituting instance.

1. Jane is either smart or kind.
2. If Justin is working on his novel, Rachel is working on her screenplay.
3. It is not the case that Tom is doing a trick and Sam is telling a joke.
4. Tom is not doing a trick, and Sam is telling a joke.
5. It is either the case that you will if you do your best, or that you won’t win.

6. Vanessa likes horses and either Jim is an astronaut or a philosopher.
7. If fish swim, they either live in fresh water or salt water.
8. If Phil's pet is either a cat or a dog, it is a mammal.
9. Some birds fly, but some do not.
10. It is not the case that all birds fly or that all pets are mammals.

Tautologies, Contradictions, and Contingent Statements

A third way of classifying a statement is by whether it is a **tautology**, a **contradiction**, or **contingent**. A statement is a tautology just in case there is no way for it to be false. A statement is a contradiction just in case there is no way for it to be true. A statement is contingent just in case there is at least one way for it to be true, and at least one way for it to be false.

How do we determine whether a statement is a tautology, a contradiction, or contingent? One way is to use a truth-table. Remember, on a truth-table, the column which is assigned truth-values last gives all the possible truth-values of the whole statement. So, if every position in that column receives the truth-value true, the statement is a tautology; if every position in that column receives the value false, the statement is a contradiction; finally, if at least one position receives the value true and at least one receives the value false, the statement is contingent.

Let us consider some examples. Begin with the statement

p & ~p.

The truth-table for this statement begins like this:

p	p & ~p
T	
F	

On the top left we have our statement letter, p, on the top right we have our statement "p & ~p," and on the bottom left we have our two possible truth-values for "p."

We want to assign truth-values to our statement, but we have two operators, ~ and &. To which do we assign truth-values first? Well, & combines p and ~p, so we cannot determine the truth-values for & until we know the truth-values for "~p." So, we begin by assigning truth-values to ~. The statement "~p" always has the opposite truth-value as "p." Thus we have:

p	p & ~p
T	F
F	T

Next, we assign truth-values to &. A conjunction is true just in case the statements on both sides, the **conjuncts**, are both true. So, we have:

p	p & ~p
T	F F
F	F T

The column showing the truth-values for the whole statement has been circled for clarity. As we can see, "p & ~p" is always false. This statement is a contradiction.

Since "~p" always has the opposite truth-value of "p," the statement "~(p & ~p)" must always be false. Here is the truth-table:

p	~(p & ~p)
T	T F F
F	T F T

As before, the column showing the truth-values for the whole statement has been circled for clarity. Hence, “~(p & ~p)” is a contradiction.

An example of a contingent statement is

(t v s) & ~(t & s).

Here is the table:

t	s	(t v s) & ~(t & s)
T	T	T F F T
T	F	T T T F
F	T	T T T F
F	F	F F T F

As we can see, the statement is true in at least one case, and false in at least one case.

Suppose that a certain statement S is a tautology. This means that S is true no matter the truth-value of its atomic parts. Suppose we make a substitution instance of S by replacing one or more of its statement letters with statements. These statements which we substitute in for the statement letters will either be contingent, contradictory, or tautologous; that is, they will either be true in some cases and false in others, always false, or always true. But, we know that S will be true regardless of whether these are true or false. So, any substitution instance of a tautology will also be a tautology. By similar reasoning we can see that any substitution instance of a contradiction will also be a contradiction.

It is not the case, however, that every substitution instance of a contingent statement is contingent. By substituting one or more of the statement letters in a contingent statement with a tautology or a contradiction, it is possible to make the whole statement necessarily true or necessarily false. For instance, if we take the contingent statement

a & b

and substitute “a v ~a” for “a,” and “b v ~b” for “b,” we get

(a v ~a) & (b v ~b),

which is a tautology. Alternatively, if we substitute “~a” for “b” we get

a & ~b,

which is a contradiction.

Work Problems 6.3

1. For each of the following statements, use a truth-table to determine whether it is a tautology, a contradiction, or contingent.

1.1. p → p

1.2. ~(p → p)

1.3. p v ~p

1.4. ~(p v ~p)

1.5. a v b

1.6. ~(a v b)

2. Explain why any substitution instance of a contradiction is a contradiction.

Argument Forms and Substitution Instances

In the previous chapter, we discussed what it means for one statement to be a substitution instance of another. We also said that a substitution instance of a statement has the same form as that statement. These concepts of form and substitution instance can be applied to arguments as well. That is, arguments have substitution instances, and every substitution instance of an argument shares a logical form with that argument.

To get a substitution instance of an argument, we just take an argument, such as

$p \rightarrow q$	(Premise)
p	(Premise)
q	(Conclusion)

and we replace any statement letter in this argument wherever it appears with whatever statement we like. Thus, if we replace p with

$p \rightarrow (q \& r)$

and we replace q with

$\sim(s \vee (t \& (u \rightarrow \sim v)))$

We get

$((p \rightarrow (q \& r)) \rightarrow \sim(s \vee (t \& (u \rightarrow \sim v)$

$p \rightarrow (q \& r)$

$\sim(s \vee (t \& (u \rightarrow \sim v)$

One form this argument has is

$p \rightarrow q$

p

q

Of course, it has another form too:

f

p

q

As with statements, an argument will share with all its substitution instances a logical form. Also, like statements, all arguments containing molecular statements are substitution instances of multiple arguments, and therefore have multiple argument forms.

Why are argument forms so important? In propositional logic, as with categorical logic, a valid argument form is an argument form every substitution instance of which is an argument which is valid. Why? An argument with a valid form is logically valid, meaning it is valid independent of the meaning of the (non-logical) terms involved. A substitution instance of an argument will have the same form, just extra content. So, if the original argument is logically valid, every substitution instance will be.

Not only is every substitution instance of a valid argument valid, every logically valid argument is also the substitution instance of some valid argument form. We see that this must be the case, because if an argument was not the substitution instance of a valid argument form, it could be only valid in virtue of the meaning of its non-logical terms, and hence at best analytically valid.

Work Problems 6.4

Give one substitution instance for each of the following argument forms:

1. a v b

   ~a

   b

2. a → b

   ~b

   ~a

3. a & b

   a

4. a ←→ b

   ~a

   ~b

Using Truth-Tables to Test for Validity

So far, truth-tables have proven very useful in our exploration of propositional logic. Truth-tables can also be used to determine whether or not an argument is (logically) valid. Remember, an argument is valid just if it cannot be the case both that all the premises are true and the conclusion is false. A truth-table gives all the possible truth-values of a statement or group of statements. So, we can determine whether an argument is valid by making a truth-table with the premises and the conclusion of the argument. If there is any row in which all the premises are true and the conclusion is false, the argument is invalid. If there are no such rows, the argument is valid. Let us consider some examples.

First, consider the argument

p → q (Premise)

p (Premise)

q (Conclusion)

We begin the truth-table like this:

p	q	p → q	p	q
T	T			
T	F			
F	T			
F	F			

Next, we assign truth-values to our statements. Since "p → q" is a conditional statement, it is true just in case the antecedent is true, or the consequent is false. The other statements "p" and "q" just retain their truth-values.

p	q	p → q	p	q
T	T	T	T	T
T	F	F	T	F
F	T	T	F	T
F	F	T	F	F

Each row shows us a possible pattern of truth-values for the premises and the conclusion. We check to see if there is any row in which all the premises are true, and the conclusion is false. There is not. So, there is no possible way for the premises of this argument to be true and for the conclusion to be false. Thus, the argument is valid. In determining that the argument was valid, we did not rely on the meaning of "p" or "q" at all. So, we know that any argument of this form will be valid. This is a valid argument form.

Let us compare a different argument:

p → q (premise)

q (premise)

p (conclusion)

Here is the table:

p	q	p → q	q	p
T	T	T	T	T
T	F	F	F	T
F	T	T	T	F
F	F	T	F	F

In the third row, which has been circled, we see that both premises are true, and the conclusion is false. So, the argument is not valid. This is not a valid argument form.

Next, consider the argument:

p → q (Premise)

q → r (Premise)

p → r (conclusion)

Here is the table:

p	q	r	p → q	q → r	p → r
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	T	T
T	F	F	F	T	F
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

Is there any row in which the premises are both true and the conclusion is false? There is not. The conclusion is false in the second row from the top and the fourth row from the top. In the second row from the top, the second premise is false. In the fourth row from the top, the first premise is false. There is no case in which both premises are true, and the conclusion is false. The argument is valid.

Compare this argument:

p → q (Premise)

p → r (Premise)

q → r (Conclusion)

Here is the table:

p	q	r	p → q	p → r	q → r
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	T	T
T	F	F	F	F	T
F	T	T	T	T	T
F	T	F	T	T	F
F	F	T	T	T	T
F	F	F	T	T	T

On the sixth row from the top, which has been circled, the premises are both true and the conclusion is false. So, the argument is invalid.

Next consider this argument:

p → q (Premise)

~q → ~p (Conclusion)

Here is the table:

p	q	p → q	~q → ~p
T	T	T	F T F
T	F	F	T F F
F	T	T	F T T
F	F	T	T T T

The column giving the truth-values for the conclusion has been circled for clarity. This argument is valid: there is no row in which the premise is true, and the conclusion is false. But notice that there is no row in which the conclusion is true, and the premise is false. So, this argument,

~q → ~p (Premise)

p → q (Conclusion)

is also valid. These two statements

p → q

and

~q → ~p

have the same truth-values in all the same cases. They are **logically equivalent**.

Work Problems 6.5

1. For each of the following arguments, use a truth-table to determine whether or not it is valid.

1.1. p

p v q

1.2. p → q

~q

~p

1.3. ~(p & q)

~p v ~q

1.4. $p \rightarrow (q \rightarrow r)$

$(p \& q) \rightarrow r$

1.5. p v (q & r)

(p v q) & (p v r)

2. For each of the following groups of three statements, use a truth-table to show that they are logically equivalent.

   2.1. a & b :: ~(~a v ~b) :: ~(a $\rightarrow$ ~b)

   2.2. a v b :: ~(~a & ~b) :: ~a $\rightarrow$ b

   2.3. a $\rightarrow$ b :: ~(a & ~b) :: ~a v b

3. Why are the logical equivalencies in part 2 significant?

4. An argument consists of at least one premise, and one conclusion. Suppose we take all the premises and **conjoin them** parentheses and as many &s as we need. Then make a conditional statement with the **conjunction** of the premises as the antecedent, and with the conclusion as the consequent. Keep this in mind when answering the following questions:

   4.1. If the original argument is valid, will the resulting statement be tautologous, contradictory, or contingent?

   4.2. If the original argument is invalid, will the resulting statement be tautologous, contradictory, or contingent?

The Counterexample Method

Previously, we saw that we could use the counterexample method to show that an argument is categorical logic does not have a valid argument form. We can use the counterexample method to show that an argument form in propositional logic is not valid as well. To do this, we just see if we can fill in the statement letters to make the premise or premises obviously true, and the conclusion obviously false. As with categorical logic, it is often useful to begin with the conclusion.

Consider the following argument form

(p & ~(q v r)) → (a ←→ b)	(Premise)
a ←→ b	(Premise)
p & ~(q v r)	(Conclusion)

To begin, we fill in the statement letters of the conclusion so that the resulting statement is obviously false. For instance, we replace p with “Dogs are reptiles”, q with “cats are mammals” and r with “birds have feathers.” Thus, our conclusion is

Dogs are reptiles and it is not the case that either cats are mammals or birds have feathers.

This statement is true only if the statements on both sides of “and” are true. Since “dogs are reptiles” is false, the whole thing is false. Since cats are mammals and birds do have feathers, the statement “either cats are mammals or birds have feathers” is true, and so the statement “It is not the case that either cats are mammals or birds have feathers” is false. So, it turns out that the statements on both sides of the “and” are false, but only one needs to be.

Next, we fill in for p, q, and r throughout the argument. The result is:

(Dogs are reptiles and it is not the case that either cats are mammals or birds have feathers) → (a ←→ b)

a ←→ b

Dogs are reptiles and it is not the case that either cats are mammals or birds have feathers.

Now we need to fill in the statement letters in the premises to make them obviously true. The first statement is a conditional, and hence is true if the antecedent is false or the consequent is true. Since the antecedent is false, it will be true no matter what we make the consequent.

The consequent of the first premise is the same as the second premise, so, we should try to make the second premise true. The second premise is a biconditional, which is true just in case both sides have the same truth-value. Let us replace a with "frogs are amphibians" and b with "snakes have forked tongues." Thus, are our argument is:

If dogs are reptiles and it is not the case that either cats are mammals or birds have feathers, then, frogs are amphibians if and only if snakes have forked tongues.

Frogs are amphibians if and only if snakes have forked tongues.

Dogs are reptiles and it is not the case that either cats are mammals or birds have feathers.

If you think about it for a little bit, you will see that the premises are both true and the conclusion is false. So, there is a way of filling in the content of our original argument form to generate an argument with true premises and a false conclusion. This means that the argument form is not valid.

So, we can use the counterexample method to show that an argument form in propositional logic is not valid. Can we use the counterexample method to show that an argument is not valid? Well, we can take an argument in a natural language, translate the statements comprising the argument into our symbolic language, and then use the counterexample method to test if the resulting argument form is valid. This alone, however, will not suffice to show that the original argument is logically invalid. Why not? Remember, every argument containing molecular statements has multiple argument forms. Even if one form is not valid, the argument may have another form that is valid.

Fortunately, for any English argument with a finite number of premises and such that the premises and conclusion are all finite in length, if it can be adequately analyzed with propositional logic, if it is invalid, we can use the counterexample method to show this.

Whenever you generate a substitution instance of a statement or argument, you replace every instance of any statement letter with any statement in the language. This means that any substitution instance of an argument form will either replace every instance of a given statement letter with a different statement letter, or with a molecular statement. If each statement letter is replaced with a different statement letter, then the very same logical operators and parentheses, in the same places, will appear in the original argument form and the substitution instance. If any of the statement letters are replaced with molecular statements, then all the logical operators and parentheses appearing in the original will appear in the substitution instance, and in the same order, but some extra operators and parentheses will appear in the substitution instance.

Given this, we can see that for any given argument we pick, for any one of its argument forms, the argument form will either have the very same number of statement letters, logical operators, and parentheses, in the same arrangement, or the argument form will have less statement letters, logical operators, and parentheses. Since we are assuming that the argument contains a finite number of statements of finite size, it can only have a finite number of argument forms. This means that in principle could use the counterexample method on every argument form of a given argument. If we can give a counterexample to every argument form, the argument must be invalid.

Even so, the more complicated an argument, the greater the number of argument forms it will have. Moreover, the more complicated a given argument form, the more effort goes into trying to come up with true premises and a false conclusion using the counterexample method. For practical reasons, then, the counterexample method is really only useful for fairly simple arguments.

Work Problems 6.6

For each of the following arguments, use the counterexample method to show it does not have a valid form.

1. If it is raining, it is cloudy.

   It is not raining.

   It is not cloudy.

2. If it is raining, it is cloudy.

   It is cloudy.

   It is rainy.

3. If you study hard, then if you go to class you will do well on the exam.

   You will not do well on the exam.

   You did not study hard.

Chapter 7. Proofs in Propositional Logic: Rules of Inference

In the previous chapter, we discussed how to use truth-tables to evaluate arguments for validity. We stuck to fairly simple arguments. As we have seen, as the number of different statement letters increases, so does the number of rows on the truth-table. An argument with 2 statement letters has four rows. An argument with 3 statement letters has 8 rows. An argument with 4 statement letters would have 16 rows. An argument with 5 statement letters would have 32 rows. Thus, as arguments become more complex, determining validity using truth-tables becomes increasingly laborious. In this chapter and the next we discuss a different method for determining validity: proofs.

We can begin by explaining the basic idea behind proofs. Shortly we will introduce a list of rules (and in the next chapter we will introduce a second list!). Each rule says that, if you have a statement or statements of the appropriate form, you may write another statement of the appropriate form. These rules are designed so that, if followed correctly, you are only allowed to write down statements which follow validly from other statements you already have.

When doing proofs, you begin with some premises and a conclusion. Your goal is to write down new statements according to the rules, and eventually to write the conclusion. Since the rules are designed so that, if followed correctly, you are only allowed to write down statements which follow validly from other statements you already have, you will be able to write the conclusion only if it follows validly from the premises. Thus, by following the rules and writing the conclusion, you will have *proved* that the argument is valid.

Here is an example. We know that the following argument form is valid:

$p \rightarrow q$ (Premise)

p (Premise)

q (Conclusion)

We know this, because we proved it on a truth-table in the previous chapter. Since this is valid, we could have a rule that says

Whenever you have a statement of the form

$p \rightarrow q$,

and a statement of the form

p,

you may write a statement of the form

q.

Now consider the following argument:

1. a → (b → c)
2. a
3. b / c

Notice that we have written the argument by numbering the premises, starting with 1, and then putting the conclusion on the last line following a slash. This is a fairly standard way of representing arguments when doing proofs.

Notice that the first premise,

a → (b → c)

is a substitution instance of

p → q,

and the second premise

a

is a substitution instance of

p.

So, the first two statements have the form

$p \rightarrow q$

p

That means we may write the appropriate statement of the form q, which would be the substitution instance of q:

$(b \rightarrow c)$.

Thus, we may write:

1. $a \rightarrow (b \rightarrow c)$
2. a
3. b
4. $b \rightarrow c$ 1, 2 rule

Next to the new line, we indicate the lines we used as premises, and the rule. Since it is the only rule we have, we simply name it "rule".

Now, notice that the new line,

$b \rightarrow c$

is a substitution instance of

$p \rightarrow q$,

and

b

on line 3 is a substitution instance of

p.

That is, lines 4 and 3 have the form

$p \rightarrow q$

p

So, we know that the substitution instance of q will follow validly from these lines. Thus, we can add a new line to our proof:

1. a $\rightarrow$ (b $\rightarrow$ c)
2. a
3. b
4. b $\rightarrow$ c 1, 2 rule
5. c 3, 4 rule

As before, we indicate the lines we used as premises, and the rule. Line 5 is c, our conclusion. Since every line we wrote down followed validly from our original premises, c must follow validly from our original premises. Thus, the original argument is valid.

That is the basic idea behind proofs. Learning to do proofs can take practice, but it is an incredibly useful way of testing arguments in propositional logic for validly. You can think of doing proofs like playing a game. You win if you can arrive at the conclusion, from the premises, using only the available rules. As with many games, the best way to learn how to play is just to play until you figure it out. To do this, though, we first need some rules.

Introducing Rules of Inference

What sort of rules should be use? In principle, we could make a rule for any valid argument form. However, it would be good for our rules to have certain features. For one, we would like for them to be fairly simple. The simpler a rule, the easier it is to remember, the easier it is to apply, and the easier it is to verify that it is in fact valid, say, by making a truth-table. For another, we want our rules to be of a manageable number. The more rules we have, the harder it will be to memorize all of them. Finally, we want enough rules that for a great many valid arguments in propositional logic, we can show that they are valid using our rules. A rule system containing only the rule used in the previous section would be easy to memorize and easy to apply, but there are lots and lots and lots of valid arguments that we could not prove are valid with that rule alone.

There are two kinds of rules for doing proofs: **Rules of inference** and **Rules of replacement**. We will talk about rules of inference in this chapter, and rules of replacement in the next. The rule used in the previous section is an example of a rule of inference. If you have a statement of the form

$p \rightarrow q$

and another statement of the form

p

you may write a new statement of the form

q.

As previously discussed, you may do this because the statement with the form q follows validly from the other two statements.

Shortly we will give a list of 8 rules of inference (we will introduce 10 rules of replacement in the next chapter). First, though, it will be good to describe the procedure by which these rules were selected. To actually go through the steps of choosing these rules here would be overly long and complicated. Still, it would be good for the reader to have a basic idea of where these rules came from.

Every statement in our symbolic language is either atomic, a negation, a conjunction, a disjunction, a conditional, or a biconditional. So, for each of these basic statement forms, we ask what follows validly from a statement of that form, and what is logically equivalent to a statement of that form.

A negation may be a negation of a conjunction, disjunction, conditional, or biconditional. So, we ask what validly follows from the negation of each of these statement forms, and what is logically equivalent to any of these statement forms.

A disjunction, conditional, or biconditional can have on either side a negation, conjunction, disjunction, conditional, or biconditional. We ask what follows validly in these cases, and what is logically equivalent to these statement forms. (We do not have to do this for conjunction, because a conjunction always entails either of its two conjuncts alone, so a conjunctive statement can always be split into its two parts, and then each can be dealt with according to its own statement form).

Next, we ask what follows validly from any pair of the foregoing statement forms. Finally, we draft a list of rules of inference and rules of replacement that allows us to prove all of these valid entailments. Here are the rules of inference (we use capital letters to indicate these are statement forms):

1. Modus Ponens P → Q P / Q MP	2. Modus Tollens P → Q ~Q / ~P MT
3. Disjunctive Syllogism P v Q ~P / Q DS	4. Hypothetical Syllogism P → Q Q → R / P → R HS
5. Addition P / P v Q Add	6. Conjunction P Q / P & Q Con
7. Simplification P & Q / P Simp	8. Constructive Dilemma (P → Q) v (R → S) P & R / Q v S CD

Each of these can be verified on a truth-table.

By the way, some of these words may be unfamiliar. That's okay. "Modus Ponens" and "Modus Tollens" are Latin names which medieval philosophers gave those argument forms.

“Modus ponens” roughly means “mode of affirmation” and “Modus tollens” roughly “mode of denying.” A “syllogism” is from a Greek word for “inference,” and refers to an argument with two or more premises.

Work Problems 7.1

Use truth-tables to prove that four of the rules of inference are valid.

Rules of Inference 1 - 4

In this section we discuss the first four rules of inference.

1. Modus Ponens P → Q P / Q MP	2. Modus Tollens P → Q ~Q / ~P MT
3. Disjunctive Syllogism P v Q ~P / Q DS	4. Hypothetical Syllogism P → Q Q → R / P → R HS

Students sometimes find proofs difficult at first. If you find them challenging, realize that you are not alone, and keep trying. This is a case in which there is really no substitute for practice. So, we are going to go through examples, beginning with very simple ones.

We can start with the following argument:

1. a
2. a → b / b

We want to try to prove that the conclusion follows from the premises. When doing proofs, the first thing you should look at is the conclusion. Your goal, ultimately, is to write that on a line, using the available rules. In this case, the premise is b. Next, you look to see if the conclusion appears anywhere in the premises. We see that it does, on line 2. So, it is a good assumption that we are going to have to use line 2 somehow. Next, we ask if there are any rules which we can apply to line 2 which will give us the conclusion. Well, we can apply modus ponens to lines 1 and 2: line 2 is a substitution instance of

$$P \rightarrow Q$$

and line 1 is a substitution instance of

$$P.$$

So, we can write down the appropriate substitution instance of Q. In this case, that is b. Hence, we may write:

1. a
2. a → b
3. b   1, 2 MP

The proof is complete. Notice that the lines to which we applied the rule modus ponens are not in the same order as in the rule. That is fine. Remember that the order of the premises does not impact the validity of an argument.

Here is a second, slightly more complicated argument:

1. ~a → b
2. a → c
3. ~c   / b

The conclusion is "b," and it appears in line 1. Is there any rule which we can apply to line 1? Not right now. Notice that if we had "~a" on a line, we could apply modus ponens to line 1. It does not matter that ~a is a negation, since

$$\sim a \rightarrow b$$

is a substitution instance of

$$P \rightarrow Q.$$

Unfortunately, we do not have "~a" on a line. So, next we look to see if a appears on any line, and we that it appears on line 2. Is there any rule which we can apply to line 2? We can apply modus tollens to lines 2 and 3, with

a → c

as our substitution instance for

P → Q

and

~c

as our substitution instance for

~Q.

So, we can write:

1. ~a → b
2. a → c
3. ~c / b
4. ~a 2, 3 MT

Now we have "~a." Recall that earlier we noted that if we had "~a" on a line by itself, we could do modus ponens with that line and line 1, and get the conclusion. So,

1. ~a → b
2. a → c
3. ~c / b
4. ~a 2, 3 MT
5. b 1, 4 MP

And the proof is complete.

Next, consider this argument

1. a → b
2. b → (c v d)
3. ~(c v d) / ~a

Does the conclusion, "~a," appear anywhere in the premises? No, but "a" appears in line 1. Is there any rule which we can apply to line 1? We can apply hypothetical syllogism to lines 1 and 2, with

a → b

as the substitution instance for

P → Q,

and

b → (c v d)

as the substitution instance for

Q → R.

Unfortunately, this will not get us "~a." How could we get "~a" from line 1? Well, if we had "~b" by itself on a line, we could do modus tollens. While "~b" does not appear on a line, "b" appears on line 2. If we could do modus tollens with line 2, we could get "~b." To do modus tollens on line 2, we need "~(c v d)" on another line. Fortunately, "~(c v d)" appears on line 3. So, we have

1. a → b
2. b → (c v d)
3. ~(c v d)   / ~a
4. ~b   2, 3 MT

This gives us our "~b," which we wanted to do modus tollens with line 1. Thus,

1. a → b
2. b → (c v d)
3. ~(c v d)   / ~a
4. ~b   2, 3 MT
5. ~a   1, 4 MT

The proof is complete. Notice that we used modus tollens twice in the same proof. That is fine. You can use a rule of inference as many times as you like, since any rule of inference will only let you write what validly follows from what you already have.

You may recall that we observed that hypothetical syllogism could be applied to lines 1 and 2. If we do that, here is what we get:

1. a → b
2. b → (c v d)
3. ~(c v d)   / ~a
4. a → (c v d)   1, 2 HS

Now we can do modus tollens with lines 3 and 4. Thus:

1. a → b
2. b → (c v d)
3. ~(c v d)   / ~a
4. a → (c v d)   1, 2 HS
5. ~a   3, 4 MT

This also completes the proof. Very often there are multiple ways of completing a proof. That is fine. You can get to the conclusion however you like, so long as you follow the rules.

Here is another argument:

1. a → (b v c)
2. ~b
3. a   / c

The conclusion is "c," which we find on line 1, where it appears as part of a disjunctive statement (a **disjunct**) which itself is the consequent of a conditional statement. Do any of our rules apply to line 1? With line 3, modus ponens applies. So,

1. a → (b v c)
2. ~b
3. a   / c

4. b v c   1, 3 MP

We drop the outer parentheses on line 4, since they are unnecessary (note that they are not unnecessary in line 1). Is there any rule which applies to line 4? Well, disjunctive syllogism with line 2, with

b v c

the substitution for

P v Q,

and

~b

the substitution instance for

~P.

Hence, we may write:

1. a → (b v c)
2. ~b
3. a   / c
4. b v c   1, 3 MP
5. c   2, 4 DS

That completes the proof.

You may notice that in each of these cases, it was crucial that we recognized when one of our rules could be applied in a given case. This requires two things: first, being able to identify substitution instances of statement forms, and second, knowing the rules. Indeed, something that will be tremendously helpful to your learning how to do proofs is to memorize the rules. Every time you look at the list of rules, you are taking your eyes, and your mind, away from the proof. Learn the rules, and you will become more efficient at doing proofs.

Work Problems 7.2

Do a proof for each of the following arguments, using only the four rules discussed so far.

1. a v (b → c)
2. ~c
3. ~a / ~b

1. f → (g v h)
2. i → f
3. i
4. ~g / h

1. a v (b v c)
2. f → ~a
3. f
4. g → ~b
5. g / c

1. (a & b) → (c v d)
2. (c v d) → ~e
3. a & b
4. e v (h → j)
5. h / j

Rules of Inference 5 – 8

Now that we have gotten some practice doing proofs with the first four rules of inference, we will add the other four. They are given on the following table:

6. Addition P / P v Q Add	7. Conjunction P Q / P & Q Con
8. Simplification P & Q / P Simp	9. Constructive Dilemma (P → Q) v (R → S) P & R / Q v S CD

Let's look at a few arguments.

We will start with this simple one:

1. e → f
2. e / f v g

The conclusion is "f v g." We find f in the first premise, but g does not appear anywhere. There is only one rule which allows us to write a statement letter (or letters) on a new line which does not appear on any previous line: addition. We can get "f v g" by adding "v g" to "f." For that purpose, we first need to get f by itself. Do any of our rules apply to line 1, which would let us get f by itself? Well, we could use modus ponens with lines 1 and 2, like so:

1. e → f
2. e / f v g
3. f 1, 2 MP

Now we use addition to get the conclusion:

1. e → f
2. e / f v g
3. f 1, 2 MP
4. f v g 3 Add

The proof is complete.

Here is another argument:

1. h → (i & j)
2. k v h
3. ~k / i

The conclusion, "i," is in the first line. It is part of a conjunction with "j" (a **conjunct**), which itself is the consequent of a conditional statement. So, we want to get the consequent by itself, and then somehow separate "i" from "j". Let's first focus on getting the consequent by itself. Is there any rule which we could apply to line 1 to get "i" & "j" by itself? Well, if we had "h" by itself on a line, we could use modus ponens. Although "h" does not appear on a line by itself, it does appear in line 2. Is there any rule which could be applied to line 2 which would let us get "h"? We can get "h" by using disjunctive syllogism on lines 2 and 3. So:

1. h → (i & j)
2. k v h
3. ~k / i
4. h 2, 3 DS

Now that we have "h" by itself, we can do modus ponens:

1. h → (i & j)
2. k v h
3. ~k / i
4. h 2, 3 DS
5. i & j 1, 4 MP

Here we have the consequent of the conditional by itself, which we wanted. All we have to do now is get "i" by itself. Is there any rule which we can apply to line 5 which will allow us to do this? Simplification:

1. h → (i & j)
2. k v h
3. ~k / i
4. h 2, 3 DS
5. i & j 1, 4 MP
6. i 5 Simp

This completes the proof.

Here is our next argument:

1. a
2. (a & b) → (c v d)
3. b
4. ~c / d

The conclusion, "d," is on line 2, where it appears as a disjunct in "c v d," which itself is a consequent of a conditional. If we could get "c v d" by itself on a line, we could use disjunctive syllogism with line 4 and get our conclusion. How could we get "c v d" by itself on a line? Well, if we had "a & b" on a line we could do modus ponens with 2. Although "a & b" does not appear on a line anywhere, "a" appears on line 1, and "b" on line 3. We can put "a" and "b" together with "&" using conjunction. Thus:

1. a
2. (a & b) → (c v d)
3. b
4. ~c / d
5. a & b 1, 3 Con

Now that we have "a & b" on line 5, we could do modus ponens with it and line 2. Hence:

1. a
2. (a & b) → (c v d)
3. b
4. ~c / d
5. a & b 1, 3 Con
6. c v d

Here we finish the proof using disjunctive syllogism on lines 4 and 6, as such:

1. a
2. (a & b) → (c v d)
3. b
4. ~c / d

5. a & b 1, 3 Con
6. c v d
7. d 4, 6 DS

This completes the proof.

Here is our next argument:

1. t
2. (t → r) v (s → u)
3. s / r v u

The conclusion, "r v u," does not appear in any of the premises, but "r" and "u" both appear in premise 2. Each is the consequent of a conditional statement, which are joined by a "v". Is there any rule which we could apply to rule 2? If we had the conjunction of the antecedents on a line, we could use constructive dilemma. Although "t & s" does not appear on a line, "t" and "s" do appear separately on lines 1 and 3, meaning that we could put them together using conjunction. Thus, we have:

1. t
2. (t → r) v (s → u)
3. s / r v u
4. s & t 1, 3 Con
5. r v u 2, 4 CD

This completes the proof.

Our final argument is:

1. f → g
2. f & h / g v j

Our conclusion is "g v j". We find "g" on line 1, but no "j". This means we are going to have to use addition. If we could get "g" by itself, we could finish the proof. We could get "g" by itself if we had "f" on a line by itself, so we could do modus ponens with line 1. While "f" does not appear on a line by itself, it does appear in line 2 conjoined with “h”. We can get “f” by itself using simplification. So, here is the proof:

1. f → g
2. f & h / g v j
3. f 2 Simp
4. g 1, 3 MP
5. g v j 4 Add

We could have also completed the proof with constructive dilemma. To do this, we would need to add “h → j” to line 1 with addition, like this:

1. f → g
2. f & h / g v j
3. (f → g) v (h → j) 1 Add

Remember, with addition we can add whatever we like, so long as we use a v. Notice, importantly, that when we added “h → j,” we needed to add parentheses to it and “f → g”. Next, we finish with constructive dilemma on lines 2 and 3:

1. f → g
2. f & h / g v j
3. (f → g) v (h → j) 1 Add
4. g v j 2, 3 CD

This completes the proof.

Work Problems 7.3

Do a proof for each of the following arguments, using only the eight rules of inference.

1. a → b
2. a & c / b v c

1. q → (r & s)

2. q & t / q & r

1. (a → b) v (c → d)
2. c & f
3. a & e
4. ~b / d

1. t
2. (t & r) → s
3. r / s v k

1. a → (b → f)
2. e → b
3. a
4. ~f / ~e v (q & r)

1. a
2. (a v b) → (~d & c)
3. d v ~g
4. h → g / ~h & a

Chapter 8. Proofs in Propositional Logic: Rules of Replacement

In this chapter, we continue our discussion of proofs in proposition logic by introducing and discussing the rules of replacement.

Introducing Rules of Replacement

Recall that a rule of inference allows us to write something on a new line, if certain things are already written on previously lines. By contrast, a rule of replacement allows us to rewrite an old line in a different way. That is, a rule of replacement says that if you have a statement or statements of the appropriate logical form, you may write down a new statement of the appropriate logical form. For instance, a rule of replacement might say that any substitution instance of

$p \rightarrow q$,

appearing by itself on a line or as part of a larger statement, may be replaced with the appropriate substitution instance of

$\sim q \rightarrow \sim p$,

and the other way; that is, any substitution instance of

$\sim q \rightarrow \sim p$

appearing by itself on a line or as part of a larger statement, may be replaced with the appropriate substitution instance of

$p \rightarrow q$.

Let us consider an example, and then explain why rules of replacement work this way.

Consider the following argument:

1. a → (b → c)
2. ~c
3. a / ~b

Notice that in the 1st premise,

b → c

is a substitution instance of

p → q.

The appropriate substitution instance for

~q → ~p

is

~c → ~b.

So, applying our new rule we may write:

1. a → (b → c)
2. ~c
3. a
4. a → (~c → ~b) 1 Rule 2

As before, we write the line or lines from which we get our new statement, along with the rule.

Then, we can use modus ponens twice, like so:

1. a → (b → c)
2. ~c
3. a
4. a → (~c → ~b) 1 Rule 2
5. ~c → ~b 3, 4 MP
6. ~b 2, 5 MP

The proof is complete.

As we have seen, a rule of replacement allows us to alter one statement inside of a larger statement. What justifies this? To answer this, let's begin by explaining why it would not be justified for a rule of replacement. Consider the following premises:

1. a → (b → c)
2. b

Why can't we use Modus Ponens to get "c" on a new line? Well, we know that if "b → c" is true, and "b" is true, then "c" must be true. According to premise 2, "b" is true. Does any premise indicate "b → c" is true? Well, "b → c" appears in premise 1, but premise 1 does not say "b → c" is true. Rather, premise 1 says "a → (b → c)" is true. But, it's possible that "a → (b → c)" is true, even if "b → c" is false. So, we don't know if "b → c" is true or not. Since we do not know if "b → c" is true, we cannot write "c" on a new line.

The situation is different with rules of replacement. In a rule of replacement, both parts of the rule are logically equivalent, meaning they are true in all the same cases. We showed that

p → q

and

~q → ~p

are logically equivalent in a previous chapter. So, "b → c" and "~c → ~b" are true in all the same cases. According to premise 1, "a → (b → c)" is true. We don't know whether "b → c" is true or not, but, we know that "b → c" and "~c → ~b" are true in all the same cases. So, since "a → (b → c)" is true, "a → (~c → ~b)" is also true, regardless of whether "b → c" is true.

Also differentiating rules of replacement from rules of inference, because both statements in a rule of replacement are logically equivalent, either can be replaced with the other. Thus, if

~b → ~a,

appears on a line, we may write

a $\rightarrow$ b

on a new line, according to our earlier rule.

Here are the ten rules of replacement:

1. Double Negation P :: ~~P DN	2. Commutativity P & Q :: Q & P Com P v Q :: Q v P Com
3. Associativity (P & Q) & R :: P & (Q & R) Assoc (P v Q) v R :: P v (Q v R) Assoc	4. Distribution P & (Q v R) :: (P & Q) v (P & R) Dist P v (Q & R) :: (P v Q) & (P v R) Dist
5. Material Implication P $\rightarrow$ Q :: ~P v Q Imp	6. Transposition P $\rightarrow$ Q :: ~Q $\rightarrow$ ~P Trans
7. DeMorgan's ~(P & Q) :: ~P v ~Q DM ~(P v Q) :: ~P & ~Q DM	8. Exportation P $\rightarrow$ (Q $\rightarrow$ R) :: (P & Q) $\rightarrow$ R Exp
9. Redundancy P :: P v P Red P :: P & P Red	10. Material Equivalence P $\leftrightarrow$ Q :: (P $\rightarrow$ Q) & (Q $\rightarrow$ P) Equiv P $\leftrightarrow$ Q :: (P & Q) v (~P & ~Q) Equiv

Each of these can be verified on a truth-table.

As with the rules of inference, some of these terms may be unfamiliar. That is okay. For the most part, these are the names that have been given to these rules of inference. The exception is redundancy. This rule is standardly called tautology, because the word "tautology" sometimes means a statement which repeats itself. However, the word "tautology" also has a technical meaning in logic: a statement which is true in all possible cases. Since this may cause confusion, we use the word "redundancy" for the rule of replacement instead. DeMorgan's is, as you may have guessed, named after a person.

Rules of Replacement 1 – 5

As with the rules of inference, we will learn the rules of replacement in two parts, beginning with rules 1 – 5, which are given on the following table:

1. Double Negation P :: ~~P DN	2. Commutativity P & Q :: Q & P Com P v Q :: Q v P Com
3. Associativity (P & Q) & R :: P & (Q & R) Assoc (P v Q) v R :: P v (Q v R) Assoc	4. Distribution P & (Q v R) :: (P & Q) v (P & R) Dist P v (Q & R) :: (P v Q) & (P v R) Dist
5. Material Implication P → Q :: ~P v Q Imp	

As with the rules of inference, we will learn these rules through practice.

Here is our first argument:

1. ~a → ~b
2. b / a

The conclusion, "a", appears negated in the first line. We would like to get "a" by itself and remove the negation. Do we have any rule which applies to line 1? Well, we could always use addition, but it is not clear how that would help us. We could use double negation on line 1, but it is not clear how that would help us either. We cannot do modus ponens with 1 and 2, because line 2 is not the antecedent of line 1. We cannot do modus tollens either, because, at least as written, line 2 is not the negation of the consequent of line 2. However, notice that since the consequent of line 1 is a negation, "~b", the negation of the consequent is "~~b". We can get "~~b" by applying double negation to line 2. So,

1. ~a → ~b
2. b / a
3. ~~b 2 DN

With "~~b", we can do modus tollens with lines 1 and 3:

1. ~a → ~b

2. b / a
3. ~~b 2 DN
4. ~~a 1, 4 MT

Here,

~a → ~b

is the substitution instance for

P → Q

and

~~b

is the substitution instance for

~P.

By double negation, “~~a” is equivalent to our conclusion, “a”. Thus:

1. ~a → ~b
2. b / a
3. ~~b 2 DN
4. ~~a 1, 4 MT
5. a 4 DN

The proof is complete.

Here is our second argument:

1. (b & c) → (d v e)
2. (e v d) → f
3. c & b / f

The conclusion, “f”, is the consequent of line 2. We could get “f” by itself if we could do modus ponens with line 2. In order to do modus ponens with 2, we need “e v d”. We see that “e v d” appears in line 2, but it is not clear how we would remove “f” to get “e v d” by itself. We also see that “d v e” appears as the consequent on line 1. So, if we could get “d v e” by itself, we could use commutativity to get “e v d”. We could get “d v e” from line 1 if we could do modus ponens on 1. For that, we would need “b & c”. Lucky for we can get “b & c” by using commutativity on line 3:

1. (b & c) → (d v e)
2. (e v d) → f
3. c & b / f
4. b & c 3 Com

Next, modus ponens with 1 and 4:

1. (b & c) → (d v e)
2. (e v d) → f
3. c & b / f
4. b & c 3 Com
5. d v e 1, 4 MP

Next, we use commutativity on 5 so we can do modus ponens with 2:

1. (b & c) → (d v e)
2. (e v d) → f
3. c & b / f
4. b & c 3 Com
5. d v e 1, 4 MP
6. e v d 5 Comm
7. f 2, 6 MP

The proof is finished.

Our third argument goes like this:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k

The conclusion, “k”, appears in line 1, as part of a conjunct with three conjunctions, which itself is the consequent of a conditional. If we could get the consequent by itself, we might be able to get “k” by itself. Is there any rule we could use to get the consequent of 1 by itself? If we had “(t v r) v s”, we could do modus ponens. We have “t v s” on line 2. We can add an “r” with addition, like so:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k
3. (t v s) v r 2 Add

Is there any way to get from “(t v s) v r” to “(t v r) v s”? We want to switch “s” and “r”. If “s” and “r” were within the same parentheses, we could switch them with commutativity. Luckily, we can move the parentheses with associativity. Thus:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k
3. (t v s) v r 2 Add
4. t v (s v r) 3 Assoc
5. t v (r v s) 4 Comm

Now we have “r” and “s” in the right order. But we didn’t want “t v (r v s)”, we wanted “(t v r) v s”. We need to move the parentheses again before we can do modus ponens. So:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k
3. (t v s) v r 2 Add
4. t v (s v r) 3 Assoc
5. t v (r v s) 4 Comm
6. (t v r) v s 5 Assoc
7. g & (h & k) 1, 6 MP

At last we have the consequent of 1 by itself. Our goal is now to get “h” by itself. If we simplify, we will get “g” by itself, which is not what we want. If we use commutativity we can get rid of “g” and keep “h & k”, and then simplify, as such:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k
3. (t v s) v r 2 Add

4. t v (s v r) 3 Assoc
5. t v (r v s) 4 Comm
6. (t v r) v s 5 Assoc
7. g & (h & k) 1, 6 MP
8. (h & k) & g 7 Com
9. h & k 8 simp

Then we can finish by commuting and simplifying again, like this:

1. ((t v r) v s) → ((g & (h & k))
2. t v s / k
3. (t v s) v r 2 Add
4. t v (s v r) 3 Assoc
5. t v (r v s) 4 Comm
6. (t v r) v s 5 Assoc
7. g & (h & k) 1, 6 MP
8. (h & k) & g 7 Com
9. h & k 8 Simp
10. k & h 9 Com
11. k 10 Simp

This completes the proof.

This is our fourth argument:

1. (a & (b v c)) → d
2. a & b / d v ~f

The conclusion, “d v ~f”, does not appear in any of the premises. However, “d” appears as the consequent in line 1. So, if we could get “d” by itself, we could get “d v ~f” using addition. We could get “d” by itself if we could do modus ponens on 1, but for that we would need the antecedent: “a & (b v c)”. We have “a & b” on line 2. We can get a “c” with addition:

1. (a & (b v c)) → d
2. a & b / d v ~f
3. (a & b) v c 2 Add

We want to move the parentheses on 3, but we cannot use associativity because that rule only applies when the operators are both the same. There is a rule which can be applied when you have & and v: distribution. First, we have to commute, like this:

1. (a & (b v c)) → d
2. a & b   / d v ~f
3. (a & b) v c   2 Add
4. c v (a & b)   3 Com
5. (c v a) & (c v b)   4 Dist

We want a & (b v c). We can get “a” by simplifying from line 2. We can get “c v b” by simplifying line 2, and then get “b v c” using commutativity. Then we can use conjunction on “a” and “c v b”, and finally do our modus ponens with line 1:

1. (a & (b v c)) → d
2. a & b   / d v ~f
3. (a & b) v c   2 Add
4. c v (a & b)   3 Com
5. (c v a) & (c v b)   4 Dist
6. a   2 Simp
7. (c v b) & (c v a)   5 Com
8. c v b   7 Simp
9. b v c   8 Com
10. a & (b v c)   6, 9 Con
11. d   1, 10 MP

At last we can derive the conclusion with addition:

1. (a & (b v c)) → d
2. a & b   / d v ~f
3. (a & b) v c   2 Add
4. c v (a & b)   3 Com
5. (c v a) & (c v b)   4 Dist
6. a   2 Simp

7. (c v b) & (c v a) 5 Com
8. c v b 7 Simp
9. b v c 8 Com
10. a & (b v c) 6, 9 Con
11. d 1, 10 MP
12. d v ~f 11 Add

The proof is finished.

Here is our fifth argument:

1. ~a v b
2. ~b v c / ~a v c

Our conclusion is "~a v c". This does not appear in any of the premises, but "~a" appears in line 1 and "c" appears in line 2. If we could get either "~a" by itself we could use addition to get "~a v c", and if we could get "c" by itself, we could use addition and commutativity to get "~a v c". Unfortunately, it is not clear how we would get either "~a" or "c" by itself. We can a disjunction from a disjunct using disjunctive syllogism, but for that we would need the negation of the other disjunct by itself, and we do not have that.

What else can we do? We could conjoin lines 1 and 2, which would give us:

(~a v b) & (~b v c).

It is not clear how to move forward from here. Simplification would just get us back to 1. We cannot use distribution because for that we would need one statement which appeared in both conjuncts. When doing a proof, sometimes it is helpful to ask what, if anything, the conclusion is equivalent to. In this case, the conclusion is equivalent to:

a → c

by material implication. Do any of our rules conclude with a conditional statement? Hypothetical syllogism does, but only if we have two conditional statement where the consequent of one is the antecedent of the other. In fact, both lines 1 and 2 are equivalent to conditional statements by material implication. Let's make them into conditional statements and see if that helps us:

1. ~a v b
2. ~b v c / ~a v c
3. a → b 1 Imp
4. b → c 2 Imp

Now we can do hypothetical syllogism on lines 3 and 4, and then material implication on the result to get the conclusion. Here it is:

1. ~a v b
2. ~b v c / ~a v c
3. a → b 1 Imp
4. b → c 2 imp
5. a → c 3, 4 HS
6. ~a v c 5 Imp

That is the proof.

Work Problems 8.1

Do a proof for each of the following arguments, using only the eight rules of inference and the four rules of replacement which we have discussed.

1. ~ ~ ~p
2. ~p → (q v (t & r))
3. ~r / q

1. (a & b) → (c v d)
2. b & a
3. ~d / c

1. (f & g) & h
2. g → ((j & k) v (j & i))
3. ~k / j

1. a
2. b v c
3. c → d / a & (~b → d)

Rules of Replacement 6 – 10

In this section we complete our discussion of the rules of replacement with rules 6 - 10:

6. Transposition P → Q :: ~Q → ~P Trans	7. DeMorgan's ~(P & Q) :: ~P v ~Q DM ~(P v Q) :: ~P & ~Q DM
8. Exportation P → (Q → R) :: (P & Q) → R Exp	9. Redundancy P :: P v P Red P :: P & P Red
10. Material Equivalence P ↔ Q :: (P → Q) & (Q → P) Equiv P ↔ Q :: (P & Q) v (~P & ~Q) Equiv	

Our first argument goes like this:

1. a → b
2. ~c → ~b / ~c → ~a

The conclusion, "~c → ~a", does not appear in any of the premises. However, "a" appears as the antecedent in line 1, and "~c" appears as the antecedent in line 2. Is there any rule which applies to both 1 and 2? Not as they are now, but maybe we can modify them so that a rule does apply. It almost looks like we could use hypothetical syllogism with lines 1 and 2, but for that we would need the consequent of one line to be the same as the antecedent of the other. We can do this with transposition on line 1, with a → b

our substitution instance of

P → Q.

Thus, we write:

1. a → b
2. ~c → ~b / ~c → ~a
3. ~b → ~a 1 Trans

Notice that the consequent of line 1 is the same as the antecedent of line 3. This means we can use hypothetical syllogism:

1. a → b
2. ~c → ~b / ~c → ~a
3. ~b → ~a 1 Trans
4. ~c → ~a 2, 3 HS

We have arrived at our conclusion.

This is our second argument:

1. ~(t & j)
2. t / ~j

The conclusion, "~j", does not appear in any of the premises, though "j" appears in line 1. Do any rules apply to line 1? Commutativity applies, but it is not so clear how it is going to help us: we want to get "j" by itself, and negated, whereas commutativity will only switch the order of "t" and "j". The rule DeMorgan's applies to 1, with

~(t & j)

being a substitution instance of

~(P & Q).

Using DeMorgan's gets us:

1. ~(t & j)

2. t / ~j
3. ~t v ~j 1 DM

Now we have our conclusion, “~j”, as a disjunct in line 3. Is there any rule that allows us to get one disjunct from a disjunctive statement? Disjunctive syllogism lets us do this, if we have the negation of the other disjunct. The other disjunct is “~t”, the negation of which is “~~t”. We can get “~~t” from line 2 with double negation, and then complete the proof with disjunctive syllogism, so like:

1. ~(t & j)
2. t / ~j
3. ~t v ~j 1 DM
4. ~~t 2 DN
5. ~j 3, 4 DS

There is the proof.

For our third argument, we will use:

1. e → (f → g)
2. (g & h) → k
3. (e & f) & h / k

Notice that the rule exportation applies to both lines 1 and 2. That gets us:

1. e → (f → g)
2. (g & h) → k
3. (e & f) & h / k
4. (e & f) → g 1 Exp
5. g → (h → k) 2 Exp

Now, we can use hypothetical syllogism on lines 4 and 5, like so:

1. e → (f → g)
2. (g & h) → k
3. (e & f) & h / k
4. (e & f) → g 1 Exp

5. g → (h → k) 2 Exp
6. (e & f) → (h → k) 4, 5 HS

If we had “e & f”, we could get “h → k” from 6 using Modus Ponens. We can get “e & f” from line 3 with simplification:

1. e → (f → g)
2. (g & h) → k
3. (e & f) & h / k
4. (e & f) → g 1 Exp
5. g → (h → k) 2 Exp
6. (e & f) → (h → k) 4, 5 HS
7. e & f 3 Simp
8. h → k 6, 7 MP

Now we could get our conclusion, “k”, with Modus Ponens on line 8 if we had “h”. Fortunately, “h” appears on line 3, and we can get it by itself with commutativity and simplification, like so:

1. e → (f → g)
2. (g & h) → k
3. (e & f) & h / k
4. (e & f) → g 1 Exp
5. g → (h → k) 2 Exp
6. (e & f) → (h → k) 4, 5 HS
7. e & f 3 Simp
8. h → k 6, 7 MP
9. h & (e & f) 3 Com
10. h 9 Simp
11. k 8, 10 MP

This completes the proof.

Our fourth argument goes like this:

1. a → s
2. ~(s & s) / ~a

The conclusion is “~a”, and “a” appears as the antecedent in line 1. We can get the negation of an antecedent using modus tollens. For that we need the negation of the consequent, or “~s”. On line 2 we have “~(s & s)”. Using commutativity will just get us “~(s & s)” again, and DeMorgan’s will get us “~s v ~s”. Fortunately, we can simply replace “s & s” in 2 with “s” using Redundancy, and then complete the proof with modus tollens, like so:

1. a → s
2. ~(s & s) / ~a
3. ~s 2 Red
4. ~a 1, 3 MT

We are done.

Here is the fifth argument:

1. ~a v b
2. ~a → ~b / a ←→ b

Our conclusion is “a ←→ b”. Since only one rule, material equivalence, lets us write a biconditional on a line, we know we will have to use that rule at some point. To get our conclusion using material equivalence, we will need one of the following statements:

(a → b) & (b → a),

or

(a & b) v (~a & ~b)

If we could get either “a & b” or “~a & ~b”, we could use addition, and maybe also commutativity, to get the second statement. It isn’t clear how we’d get either “a & b” or “~a & ~b”. So, let’s try for the first statement: “(a → b) & (b → a)”. Line 1 is equivalent to “a →” b by material implication, and line 2 is equivalent to “b →” a by transposition. Then we can use conjunction to combine the results, and finish with material equivalence. Here goes:

1. ~a v b

2. ~a → ~b / a ↔ b
3. a → b 1 Imp
4. b → a 2 Trans
5. (a → b) & (b → a) 3, 4 Con
6. a ↔ b 5 Equiv

This completes the proof.

Work Problems 8.2

Do a proof for each of the following arguments. You may use any of the rules of replacement or rules of inference.

1. a / (a v b) & (a v c)

1. e & g / e ↔ g

1. ~(~a v ~b)
2. (b & c) → d
3. c / d

1. (a → b) → (c → d)
2. ~ ~(a → b)
3. ~d / ~c

1. (r v r) → (p → q)
2. r & p / q

1. (a & b) v (a & d)
2. (b v c) → ~a / ~b

Chapter 9. Proofs in Propositional Logic: Conditional and Indirect Proof

In this chapter, we introduce two strategies for doing proofs in propositional logic: **conditional proof** and **indirect proof**. Before introducing conditional and indirect proofs, we first explain why they are needed.

Limitations of the Rules

By now we have gotten a lot of practice doing proofs. Consider the following argument:

1. a → b
2. (b & a) → e / a → e

Can we prove, using our rules, that the conclusion follows validly from the premises? Well, the conclusion, a → e, is a conditional statement. The antecedent is the antecedent of line 1, and the consequent is the consequent of line 2. This suggests that we should use hypothetical syllogism. To do that, however, the consequent of line 1 must be the same as the antecedent of line 2, and they are not. The consequent of 1 is b, and the antecedent of 2 is b & a. We could do hypothetical syllogism if we could go from

a → b

to

a → (b & a),

or from

(b & a) → e

to

b → e.

Is there any rule which lets us do this? It does not seem like it. We cannot use conjunction on line 1 to get a → (b & a), first because we do not have a by itself, and second because if we did we would get (a → b) & a, and there is no rule which allows to move those parentheses. Using addition on line 1 will get us a statement with v, which is also not what we want. We cannot simply on line 2 to get b → e, because line 2 is not a conjunction. It seems like we cannot prove the conclusion follows from the premises.

Should we conclude, then, that the argument is invalid? Well, let's make a truth-table and see:

a	b	e	a → b	(b & a)	→ e	a → e
T	T	T	T	T	T	T
T	T	F	T	T	F	F
T	F	T	F	F	T	T
T	F	F	F	F	T	F
F	T	T	T	F	T	T
F	T	F	T	F	T	T
F	F	T	T	F	T	T
F	F	F	T	F	T	T

Are there any rows on the table in which both premises are true, and the conclusion is false? Both premises are true on the 1st, 5th, 6th, 7th, and 8th rows from the top. The conclusion is true on each of those rows as well. So, the argument is valid.

It would seem that we have a valid argument which cannot be proven valid using our rules. Of course, every substitution instance of this argument form will also be a valid argument which cannot be proven valid using our rules. We may conclude that there are in principle infinitely many valid arguments which cannot be proven valid using our rules.

We want to expand our proof system so that we can prove a wider range of valid arguments to be valid. In order to do this, we will not so much add new rules, as new strategies or techniques for doing proofs, though one's which would not have been allowed previously. These are **conditional proof** and **indirect proofs**.

Conditional Proof

Conditional proof is a technique for proving that a conditional statement is true. We will explain how to use conditional proof using our previous argument as an example. Here it is:

1. a → b
2. (b & a) → e / a → e

The goal of conditional proof is to show that a conditional statement follows validly from some premise or premises. The first step of conditional proof is to assume the antecedent of the conditional you want to prove. You write the antecedent, in this case a, on an indented line, with ACP for **assumed for conditional proof**, like this:

1. a → b
2. (b & a) → e / a → e
    3. a ACP

Next you prove the consequent of the conditional, using the rules. For instance:

1. a → b
2. (b & a) → e / a → e
    3. a ACP
    4. b 1, 3 MP
    5. b & a 3, 4 Con
    6. e 2, 5 MP

Once you have derived the consequent, you may write the conditional on a new, unintended line, assign it the numbers of the line on which the antecedent was assumed, and all the lines used to derive the consequent, and then write CP for **conditional proof**, like so:

1. a → b

2. (b & a) → e   / a → e

   3. a   ACP

   4. b   1, 3 MP

   5. b & a   3, 4 Con

   6. e   2, 5 MP

7. a → e   3-6 CP

The proof is complete.

We want to be sure that when using conditional proof, we will only infer statements which follow validly from our premises. Here is a justification for conditional proof. An argument is valid just if it cannot be both the case that the premises are all true and the conclusion is false. A conditional statement is true just in case either the antecedent is false or the consequent is true. Suppose that in every case in which the antecedent and the other premises are true, then the consequent is true. In that case, the only other possibility when the other premises are true is that the antecedent is false, and if the antecedent is false, the conditional is automatically true. So, if in every case in which the antecedent and the other premises are true, the consequent is true, it must be the case that in every case in which the premises are true, the conditional is true. Thus, it must be the case that the conditional follows validly from the premises. So, we assume the antecedent to be true, and we see if we can derive the consequent from the antecedent and the other premises using only rules which we know to be valid. If we can do this, the conditional must follow validly from the premises.

Let's look at another argument:

1. b
2. c   / (a → b) → (a → c)

Again, we will use conditional proof. What should we assume for our antecedent? Let's try a → b, like so:

1. b
2. c / (a → b) → (a → c)
    3. a → b ACP

Now what do we do? We want to get a → c. To do that, we need to get a by itself. How could we do that? Well, we could use conditional proof again, like this:

1. b
2. c / (a → b) → (a → c)
    3. a → b ACP
        4. a ACP

Notice that we have made an assumption for conditional proof inside another assumption for conditional proof. That is fine. As long as we follow the procedure for conditional proof correctly, you may make as many assumptions for conditional proof as you like. Now we want to get c by itself. That is easy enough using redundancy from line 2, and simplification:

1. b
2. c / (a → b) → (a → c)
    3. a → b ACP
        4. a ACP
        5. c & c 2 Red
        6. c 5 Simp

Now that we have c, the consequent of our second conditional, we can close out our second conditional proof:

1. b
2. c / (a → b) → (a → c)
    3. a → b ACP
        4. a ACP
        5. c & c 2 Red
        6. c 5 Simp
    7. a → c 4 – 6 CP

Notice that a → c is the consequent of our conclusion. So, we can close out the first conditional proof:

1. b		
2. c	/ (a → b) → (a → c)	
3. a → b		ACP
	4. a	ACP
	5. c & c	2 Red
	6. c	5 Simp
7. a → c		4 – 6 CP
8. (a → b) → (a → c)		3-7 CP

This completes the proof.

Earlier we showed that there are valid arguments that cannot be proven valid using just our 8 rules of inference and 10 rules of replacement. You may wonder if, instead of introducing conditional proof, we could have added more rules like the old to deal with these arguments? Well, perhaps you could have made rules for this purpose, but there is a further reason we need conditional proof.

Recall that a statement is a **tautology** just in case it is true in all possible cases. Since a valid argument is one such that it cannot both be the case that the premises are all true and the conclusion false, and since there is no way for a tautological statement to be false, any argument with a tautology for its conclusion is automatically valid. That means we should want to be able to prove valid any argument with a tautology as its conclusion, regardless of what the premises are; indeed, we should want to be able to prove these statements true without requiring any premises at all. However, all of our rules of inference and replacement allow us to write new lines based on lines we already have, and this would be true of any other rules like them.

Consider the statement a → a. Either a is true or a is false. If a is true, the consequent is true, and so the statement is true. If a is false, the antecedent is false and so the statement is true. So, whether a is true or a is false, a → a is true. Thus, a → a is a tautology. So, we should be able to prove it without relying on any premise whatsoever. Here goes:

1. / a → a	
2. a	ACP
3. a v a	2 Add
4. a	3 Red
5. a → a	2-4 CP

That is the proof.

Next consider this statement: ~(b & ~b). Now, clearly b & ~b is false. It is a conjunction, so it will only be true if both conjuncts are true. The negation of b will always have the opposite truth-value as b. So, exactly one of b and ~b will be false, so that b & ~b will be false. But if b & ~b is false, its negation will have the opposite truth-value. Thus, ~(b & ~b) must be true. So, ~(b & ~b) is a tautology. That means we should be able to prove it. Unfortunately,

~(b & ~b)

is not a conditional!

Fortunately, we can prove it, we just need to be strategic. We need to find a conditional statement which is equivalent to this statement. Then we can prove that conditional statement, and derive our conclusion using our rules of replacement. Here is how we might proceed:

1. / ~(b & ~b)	
2. b	ACP
3. b v b	2 Add
4. b	3 Red
5. b → b	2-4 CP
6. ~b v b	5 Imp

7. ~~(~b v b) 6 DN
8. ~(~~b & ~b) 7 DN
9. ~(b & ~b) 8 DN

That is the proof. This is a good example of the fact that sometimes when solving a problem using conditional proof, you do not use conditional proof to get to the conclusion; instead, you use conditional proof to get to something you need to get to the conclusion.

Indirect Proof

The other technique for doing proofs is **indirect proof**. As with conditional proof, we will explain indirect proof through an example:

1. f / ~f → g

When doing indirect proof, we begin by putting the negation of what we want on an indented line, with AIP for **assumed for indirect proof**. In this case, we will assume the negation of our conclusion:

1. f / ~f → g
    2. ~(~f → g) AIP

Then we proceed to use any of the 8 rules of inference and 10 rules of replacement to derive a contradiction: a statement of the form P & ~P. Like this:

1. f / ~f → g
    2. ~(~f → g) AIP
    3. ~(~~f v g) 2 Imp
    4. ~(f v g) 3 DN
    5. ~f & ~g 4 DM
    6. ~f 5 Simp
    7. f & ~f 1, 6 Con

Once we have derived a contradiction, on the next unintended line we may write the negation of the statement assumed for indirect proof, followed by the number of the line on which the assumption for indirect proof was made and all the lines used to derive the contradiction, followed by IP, for **indirect proof**. Thus:

1. f		/ ~f → g
	2. ~(~f → g)	AIP
	3. ~(~~f v g)	2 Imp
	4. ~(f v g)	3 DN
	5. ~f & ~g	4 DM
	6. ~f	5 Simp
	7. f & ~f	1, 6 Con
8. ~~(~f → g)		2 – 7 IP

In this case, all we need to do to get our conclusion is use double negation. So:

1. f		/ ~f → g
	2. ~(~f → g)	AIP
	3. ~(~~f v g)	2 Imp
	4. ~(f v g)	3 DN
	5. ~f & ~g	4 DM
	6. ~f	5 Simp
	7. f & ~f	1, 6 Con
8. ~~(~f → g)		2 – 7 IP
9. ~f → g		8 DN

This completes the proof.

As with conditional proof, we want to know whether indirect proof only lets us prove something if it follows validly from our premises. Here is a justification for indirect proof. An argument is valid just if there is no way for the premises to all be true and for the conclusion to be false. This means that if an argument is valid, by assuming that the premises are all true and the conclusion is false, we should be able to derive a contradiction. If the argument is invalid, there should be a way for the conclusion to be false when all the premises are true, which means that we should not be able to derive a contradiction from the assumption that the conclusion is false. Thus, if from a group of premises, we can derive a contradiction by assuming another statement, we know that the negation of that statement must follow validly from the premises.

Just as we can use multiple instances of conditional proof within the same proof, we can use multiple instances of indirect proof. Furthermore, we can even use conditional and indirect proof together. Consider the following argument:

1. r / s → (~r → t)

To prove this, we will assume the antecedent, s, using conditional proof, and then the consequent ~r → t, using indirect proof. Here it is:

1. r / s → (~r → t)
2. s ACP
3. ~(~r → t) AIP
4. ~(~~r v t) 3 Imp
5. ~(r v t) 4 DN
6. ~r & ~t 5 DM
7. ~r 6 Simp
8. r & ~r 1, 7 Con
9. ~~(~r → t) 3 – 8 IP
10. ~r → t 9 DN
11. s → (~r → t) 2 – 10 CP

The proof is complete.

Work Problems 9.1

Do a proof for each of the following arguments, using any of the rules of inference and replacement, as well as conditional and/or indirect proof.

1. (x v y) → z
2. (t v s) → (w & x) / s → z

1. / a → (b → a)

Part III: Inductive Arguments

Chapter 10. Statistical Syllogisms and Probability Theory

In chapters 3 – 9 we looked at deductive arguments. In the next several chapters we will turn our attention to inductive arguments. Recall that an inductive argument is one in which the premises either make the conclusion likely, or the argument fails. An inductive argument in which it is unlikely for the premises to be true and the conclusion false, is said to be **strong**. An inductive argument which is not strong is **weak**. A strong inductive argument with all true premises is **cogent**.

The reason it is important to learn about inductive arguments is that we are often presented with reasons to believe a claim which do not establish the truth of that claim conclusively. As such, an argument with those reasons stated as premises and the claim as the conclusion would not be deductively valid, but it might be inductively strong. If it is also cogent, then the conclusion is likely to be true.

Thus, inductive arguments can give us good reasons to believe certain claims, and so we want to know how to evaluate them.

In this chapter will be looking at a kind of inductive argument called **statistical syllogism**. A statistical syllogism is:

> *An argument which proceeds from a claim about a population to a claim about some individual or individuals within that population*.

For example, consider the following argument:

> *Dr. Monica Juarez must like Indian food, because she's a professor at College University, and 80% of professors at College University like Indian cuisine*.

Here it is again, with the premises and conclusion specified:

80% of professors at College University like Indian cuisine. (Premise)

Dr. Monica Juarez is a professor at College University. (Premise)

Dr. Monica Juarez likes Indian food. (Conclusion)

One of the premises states a fact about a population: that 80% of professors at College University like Indian cuisine. The conclusion makes a claim about some member of the population: that Dr. Monica Juarez likes Indian food.

This is a nice argument, since it is clearly inductively strong. It is of course possible that Dr. Juarez is among the 20% of professors at College University who do not care for Indian food. But, if all we know about Dr. Juarez is that she is a professor at College University, then supposing we have good reason to believe that 80% of the professor's there like Indian food, we have good reason to believe Dr. Juarez likes Indian food.

However, consider this argument:

Dr. Monica Juarez and Dr. Jamal Smith must like Indian food, because both are professors at College University, and 80% of professors at College University like Indian cuisine.

Here it is again, with the premises and conclusion specified:

80% of professors at College University like Indian cuisine. (Premise)

Dr. Monica Juarez and Dr. Jamal Smith are professors at College University. (Premise)

Dr. Monica Juarez and Dr. Jamal Smith like Indian food. (Conclusion)

While it is obvious that the premises provide some support for the conclusion, it may not be obvious how much. We can evaluate more complex statistical syllogisms like this one using the tools of probability theory, which will be the topic of the remainder of this chapter.

What are Probabilities?

Recall our previous argument:

Dr. Monica Juarez must like Indian food, because she's a professor at College University, and 80% of professors at College University like Indian cuisine.

Among the premises is an assertion about a certain proportion of professors at College University; specifically, that 80% of them like Indian cuisine. On that basis, we assign a certain probability, 80%, to the conclusion: Dr. Monica Juarez likes Indian food. But, what accounts for the connection between the proportion in the premises and the probability assigned to the conclusion? The answer to this question sets the stage for our discussion of probability theory, and how probability theory can be used to evaluate statistical syllogisms of greater complexity.

Statistical syllogisms contain claims about probability. We can represent the probability of A by a number n between 0 and 1, that is

$P(A) = 0 \leq n \leq 1$,

where

$P(A) = 0$

means that A is as unlikely (as improbable) as something could be, and

$P(A) = 1$

means that A is as likely (as probable) as something could be. But, what is being measured? That is, what do we mean by a probability?

Suppose that a standard deck of 52 cards has been shuffled thoroughly and placed face-down in front of you. What is the probability that the top card is, say, the 2 of diamonds? Two different answers suggest themselves. First, you might say that the probability that the top card is the 2 of diamonds is 1 out of 52, on the grounds that each card has an equal chance of being on top. Thus:

P(the top card is 2 of diamonds) = 1/52.

On the other hand, you might say that the probability of the top card being the 2 of diamonds is either 0 or 1, on the grounds that the card on top either is the 2 of diamonds or it is not. If it is the 2 of diamonds,

P(the top card is the 2 of diamonds) = 1,

and if it is not the 2 of diamonds,

P(the top card is the 2 of diamonds) = 0.

Which answer is correct?

Well, both are correct answers to different questions. The **objective probability** that the top card is the 2 of diamonds is 1 or 0. As a matter of objective fact, the card is either the two of diamonds or it is not, and that is that. On the other hand, the **epistemic probability** that the top card is the 2 of diamonds is 1 out of 52, since there is one 2 of diamonds in the deck, there are a total of 52 cards, and we have no reason to think any particular card is on top. Objective probability concerns what is the case, whereas epistemic probability concerns what it is reasonable to believe. Since we are interested in what it is reasonable to believe, our concern in this chapter is with epistemic probability.

Before turning to epistemic probability, however, consider whether it is possible for the objective probability of an event to be something other than 0 or 1. This could be the case if there is a process which is objectively deterministic, meaning that the total causal factors contributing to the process do not wholly fix what the outcome of the process will be.

Imagine that a certain coin has the property that there is objectively a 70% chance it will land heads when flipped. Consider the following question:

If the coin is flipped twice, what is the probability that it lands heads on both throws?

Does the word "probability" here refer to objective probability or epistemic probability?

Suppose that we have thoroughly studied the coin, and on the basis of our examination have discovered that there is objectively a 70% chance it will land heads when flipped. This being the case, it would be reasonable to be 70% confidence that the coin will land head when flipped. In that case, the objective probability and the epistemic probability will be the same.

Alternatively, suppose we have not studied the coin, and have no evidence of its peculiar nature. In that case, we would have no good basis of assigning a 70% probability to it landing heads when flipped. Since in that situation we would have no more reason to think the coin will land heads than tails, we should assign the probability that it lands heads ½, or 50%, which is again the epistemic probability.

Thus, in either case, we can understand the question to be about epistemic probability. This is because either the objective probability and the epistemic probability are the same, or they are different, but we must use the epistemic probability because the objective probability is unknown. This means that we can safely focus our attention just on epistemic probabilities.

The epistemic probability of an event or state is a measure of how reasonable it is for you to believe the event will occur or the state will obtain. But how is epistemic probability to be measured? Suppose you have a standard, thoroughly shuffled deck of 52 cards. You have no good reason to suspect any card in particular as being on top. Since a standard deck has 12 face cards (four jacks, four queens, and four kings),

P(the top card is a face) = 12/52 = 3/13.

Thus, in this case the epistemic probability that the top card is a face is equal to the proportion of face cards tot total cards in the deck. Generalizing, it would seem that the probability of A is the number of possible A cases over the number of total cases. Unfortunately, however, this approach to measuring epistemic probability will not work in all cases.

For one thing, often when dealing with probabilities we do not consider all possible outcomes, and instead restrict our attention to certain relevant possibilities. What is the probability that a fair coin flipped will land either heads or tails? We want to answer

P(fair coin lands heads or tails) = 1.

However, it is possible, if incredibly unlikely, that the coin would land on its edge. Similarly, it is conceivable that when you draw the top card from the deck, the ink mysteriously evaporates, so that upon turning the card over you find it does not count as any card in a standard deck. We typically ignore possibilities like these.

For a second problem with this approach to measuring epistemic probability, consider the probability that a certain fair coin lands heads when flipped. In this scenario, the total A cases would presumably be the total number of flips on which the coin lands heads. What is the total number of cases? Perhaps it is the number of times the coin is flipped, whether it lands tails or not. But, there will may be no record of every flip. Furthermore, even if the coin is fair, it may just so happen by chance that it lands tails more times than heads. Hence, this way of measuring epistemic probability does not necessarily get the correct result that if the coin is fair, the probability that it lands heads is ½.

Maybe the probability that a coin lands heads should be treated as the number of ways it can land heads (one) over the total number of ways it can possibly land (2). But this will not work either. Consider a case in which you have a weighted coin, so that when it is flipped it lands heads two thirds of the time (and we know this). What is the probability that the next time it is flipped, the coin will and heads? Well, there are two possible ways for the coin to land (ignoring the possibility that it could land on its side), and one of these is heads. So, if the probability of A is the number of possible outcomes in which A is the case over the total number of (relevant) possible outcomes, then

P(the coin lands heads) = ½.

But this does not seem right, for surely

P(the coin lands heads) = 2/3.

It would thus seem that our proposal is incorrect.

Instead, we should understand the probability of A as what we would reasonably expect the proportion of A cases out of all cases would tend to if a certain event were to occur or procedure to be performed an indefinite number of times. As the number of times you flip a fair coin increases, we would reasonably expect the proportion of occasions on which the coin lands heads to approach ½. If you were to shuffle a standard deck of 52 cards and draw the top card, and over again, you would reasonably expect the proportion of times you drew a face to approach 12/52 over time.

Of course, it is conceivable that you flip a fair coin again and again and it keeps landing heads without cessation, though this would be incredibly unlikely. This is the reason we talk about the proportion of x cases to total cases we would *reasonably expect*. This raises the question: what proportion of x cases to total cases should we reasonably expect? In practice we have to make the best estimate we can using our background information. We know from experience that most coins are at least very close to fair, that a newly opened deck of cards (with any instructions and jokers removed) has 52 members, and so on.

Calculating Probabilities

Recall the following statistical syllogism:

Dr. Monica Juarez and Dr. Jamal Smith must like Indian food, because both are professors at College University, and 80% of professors at College University like Indian cuisine.

Earlier, we said that while the premises obviously provide some support for the conclusion, it is not clear how much. Now we can see that the amount of support the premises provide for the conclusion is just the probability that both Dr. Juarez and Dr. Smith like Indian food. Probability theory provides us with a means for determining that value. In the next several sections we will discuss how to calculate complex probabilities from simple ones.

Calculating the Probability that an Event Does Not Occur

We begin by discussing how to calculate the probability that something does not occur. For instance, if I flip a coin, what is the probability that it does not land heads? The answer is all the ways the coin could land except for landing heads. Ignoring the possibility that the coin land on its side, there is one way the coin can land other than landing heads: it could land tails. As another example, assume that a certain jar has 3 marbles: 2 blue and 1 red. If you draw a marble, the probability that you do not draw a blue marble is just the probability that you draw a red marble.

In general, the probability that A does not occur is just the probability of all relevant outcomes other than A. We assume that at least one of the relevant possible outcomes will occur.

This is become there must be some outcome, and we are ignoring special possibilities which we set aside (such as a coin landing on its side). Earlier we said that we assign

P = 1

to whatever is as probable as can be. So,

P(At least one of the relevant possible outcomes occurs) = 1.

So, the probability that A does not occur, that is, the probability that some outcome other than A occurs, is just 1 minus the probability that A occurs. That is,

P(Not A) = 1 – P(A).

Using our examples:

P(Not land heads) = 1 – P(Land heads) = 1 – ½ = ½,

and

P(Not draw blue) = 1 – P(Draw blue) = 1 – 2/3 = 1/3.

Calculating the Probability that Two Events Both Occur

Next, we will discuss how to calculate the probability that two events both occur. Suppose that a certain jar has 20 total marbles, 10 red and 10 blue. You draw two marbles in a row from the jar, without replacing the first marble before drawing the second. What is the probability that you draw a red marble both times?

To start, what is the probability that you get a red marble on the first draw? Well, clearly

P(red marble on 1^{st} draw) = 10/20 = ½.

Supposing you in fact get a red marble on the first draw, what is the probability that you get another red marble on the second draw? The answer is

P(red marble on 2^{nd} draw) = 9/19,

since 9 out of the 19 remaining marbles are red.

So then, what is the probability that you get a red marble on both draws? Well, it is reasonable to believe that if you were to repeat the task over and over again, on the first draw you would get a red marble 10/20 times. In any case in which you get a red marble on the first draw, it is reasonable to believe if you were to repeat the task over and over again, you would get another red marble on the second draw 9/19 times. So, the probability that you draw a red marble both times is the 9 out of 19 cases that you draw the second red marble for each of 10 out of 10 cases in which you get red on the first draw. That is,

P(Draw a red marble both times) = 10/20 x 9/19 = 90/380 = 9/38.

We don't just want to know the probability that you draw two red marbles in a row, but the probability of two events generally.

Here is the formula:

P(A & B) = P(A) x P(B|A)

"P(B|A)" is read "The probability of B given A." It stands for the probability that B occurs, assuming that A occurred. Here is how the formula looks when applied to our previous example:

P(Red marble on 1^{st} draw & Red marble on 2^{nd} draw) = P(Red marble on 1^{st} draw) x P(Red marble on 2^{nd} draw | Red marble on 2^{nd} draw).

Thus,

= 10/20 x 9/20 = 90/380 = 9/38,

just as we expect.

Alternatively, suppose you have two jars with 20 marbles each, 10 red and 10 blue, and you draw one marble from each jar. What is the probability that both marbles you draw are red? Well, P(Red from 1^{st} jar) = ½,

and

P(Red from 2^{nd} jar) = ½.

But what is the probability that you draw a red marble from the second jar given that you draw a red marble from the first jar? Since the jars are separate, what you draw from the first jar will have no impact on what you draw from the second jar. So,

P(Red from 2^{nd} jar | Red from 1^{st} jar) = P(Red from 2^{nd} jar) = ½.

Hence, in this case,

P(Red from 1^{st} jar & Red from 2^{nd} jar) = P(Red from 1^{st} jar) x P(Red from 2^{nd} jar | Red from 1^{st} jar)

= P(Red from 1^{st} jar) x P(Red from 2^{nd} jar) = ½ x ½ = ¼.

If A and B are independent, the probability of B given A will be the same as the probability of B.

What if we want to calculate the probability that each of many possible events occur? For instance, what is the probability that you draw 10 red marbles in a row from a jar with 10 red and 10 blue marbles? Our formula is

P(A & B) = P(A) x P(B|A).

Drawing 10 red marbles in a row is the same as drawing 9 marbles in a row and also drawing the tenth marble. So,

P(Red on 9 draws & Red on 10^{th} draw) = P(Red on 9 draws) x P(Red on 10^{th} draw | Red on 9 draws).

If you get a red marble on each of the first 9 draws, there will be one red marble remaining out of 11 marbles, so

P(Red on 10^{th} draw | Red on 9 draws) = 1/11.

Hence,

P(Red on 9 draws & Red on 10^{th} draw) = P(Red on 9 draws) x 1/11.

Now we just need to find the value for

P(Red on 9 draws).

Clearly,

P(Red on 9 draws) = P(Red on 8 draws) x P(Red on 9^{th} draw | Red on 8 draws).

Just as well,

P(Red on 8 draws) = P(Red on 7 draws) x P(Red on 8^{th} draw | Red on 7 draws),

and so on. Putting this together we get:

P(Red on 10 draws) = P(Red on 1^{st} draw) x P(Red on 2^{nd} | Red on 1^{st}) x P(Red on 3^{rd} draw | Red on 1^{st} & 2^{nd}) x … x P(Red on 10^{th} | Red on 1^{st} through 9^{th} draws),

or

10/20 x 9/19 x 8/18 x 7/17 x 6/16 x 5/15 x 4/14 x 3/13 x 2/12 x 1/11.

Calculating the Probability of Either of Two Events Occurs

Now we turn to calculating the probability that either of two different events occur, we will discuss how to calculate the probability that either of two different events occur. For example, we might wonder what is the probability that from a standard deck of 52 cards, you draw either a red card or an even card?

A standard deck contains 26 red cards and 26 black cards. So,

P(Draw a red card) = 26/52

A standard deck contains 20 even cards: four each of 2, 4, 6, 8, and 10. So,

P(Draw an even card) = 20/52.

Were you to shuffle and draw 52 times, you can reasonably expect that 26 of those times you would draw a red card, and that 20 of those times you would draw an even card.

Were you to shuffle and draw over and over again, you can reasonably expect that 26 out of 52 times you would draw a red card. Since if you draw a red card you draw either a red or an even card, each of these 26 out of 52 draws you would count as a success. Were you to shuffle and draw over and over again, you can reasonable expect that 20 out of 52 times you would draw an even card. Since if you draw an even card you draw either a red or an even card, each of those 20 out of 52 draws you would count as a success.

On some of those successful draws, however, you will have drawn a card which is both red and even. If you count ever occasion on which you draw a red card as a success, and each occasion on which you draw an even card as a success, then you will count the times during which you draw a card which is both red and even twice. Obviously, however, each of these cases only constitutes a single case of drawing either a red or even card.

This means that, from the total number of cases in which you draw a red card and the total number of cases in which you draw an even card, you want to remove the number of cases in which you draw a card which is both even and red. This prevents this case from being counted as a successful draw twice-over.

Putting this together, our formula is:

P(A v B) = P(A) + P(B) – P(A & B).

Applying this formula to our current example yields the following:

P(Draw a red or an even card) = P(Draw a red card) + P(Draw an even card) – P(Draw a red and even card)

= 26/52 + 20/52 – 10/52 = 46/52 – 10/52 = 46/52 = 11/13.

It is very likely that you will draw either a red or an even card.

What is the probability that you draw either a red or a black card? Using our formula, we have:

P(Draw red or black card) = P(Draw red card) + P(Draw black card) – P(Draw red and black card).

Since no card is both red and black, the probability that you draw such a card is zero. So,

P(Draw red or black card) = ½ + ½ - 0 = 1

When calculating the probability that either of two events occurs, it is crucial to know whether or not those events can occur together.

What if we want to calculate the probability that any one of many outcomes occurs? Well, we find the total number of successful outcomes, and subtract the total number of cases which are counted multiple times. For instance, suppose I want to calculate the probability that I draw either a red card, an even card, or an odd card. Drawing a red card, an even card, and an odd card all count as successful draws. However, I do not want to count any successful draws more than once. That means I want to exclude subtract from my total successes all those involving a card which is both red and even, red and odd, even and odd, or all three.

Thus, we have the following:

P(Draw a red, even, or odd card) = P(Red) + P(Even) + P(Odd) – (P(Red and Even) + P(Red and Odd) + (Even and Odd) + (Red, Even, and Odd)).

No card can be both even and odd, nor can any card be red, even, and odd. So,

P(Red) + P(Even) + P(Odd) – (P(Red and even) + P(Red and Odd)).

= ½ + 20/52 + 20/52 – (10/52 + 10/52)

= 26/52 + 20/52 + 20/52 – (20/52) = 66/52 – 20/52 = 46/52.

Work Problems 10.1

Suppose a jar contains 5 blue marbles, 6 red marbles, and 4 yellow marbles. Answer the following questions:

1. What is the probability you draw either a blue or a red marble?
2. What is the probability you draw either a blue or a yellow marble?
3. What is the probability you draw either a red or a yellow marble?
4. What is the probability that you do not draw a blue marble?
5. What is the probability that you do not draw a red marble?
6. What is the probability that you do not draw either a blue or a red marble?
7. What is the probability that you do not draw either a blue or a yellow marble?
8. What is the probability that you draw a blue marble, replace it, and then draw a red marble?
9. What is the probability that you draw a blue marble, replace it, and then draw a yellow marble?
10. What is the probability that you draw a blue marble, and then without replacing it draw a red marble?
11. What is the probability that you draw a red marble, and then without replacing it draw a blue marble?
12. What is the probability of getting at least one blue marble in two draws?
13. What is the probability of getting at least one red marble in two draws?

Bayes' Theorem

The last rule for calculating probabilities we will discuss is Bayes' theorem, named after Thomas Bayes. We will begin by deriving Bayes' theorem from our rule for calculating the probability of both of two outcomes. Recall,

P(A & B) = P(A) x P(B|A).

Clearly,

P(A & B) = P(B & A),

so,

P(A) x P(B|A) = P(B) x P(A|B).

Dividing each side by P(A) gets us:

$$\boldsymbol{P(B|A) = \frac{P(B) \times P(A|B)}{P(A)}}$$

This is Bayes' theorem.

To understand the significance of Bayes' theorem, it may be useful to replace B and A with something more specific. Where H is some hypothesis and E is some evidence, Bayes' theorem says:

$$\boldsymbol{P(H|E) = \frac{P(H) \times P(E|H)}{P(E)}}$$

Say we obtain some evidence E, and we want to determine how likely a hypothesis H is given E. Bayes' theorem allows us to answer this question, provided that we know:

P(H): The likelihood of the hypothesis being correct prior to the evidence.

P(E|H): The likelihood of the evidence given that the hypothesis is true.

P(E): The likelihood of the evidence independent of the hypothesis.

Thus, Bayes' theorem tells us how to reevaluate how likely something is to be the case based on new evidence.

Suppose that it is discovered that 15% of the people in your community have contracted and disease. Wishing to know whether you have the disease, you take the test and the results come back positive. How likely is it that you in fact gave the disease, given the positive test results? We can answer this question using Bayes' theorem if we know the accuracy of the test.

Specifically, we want to know in what percentage of cases the test yields a positive result when a person with the disease is tested, and we want to know what percentage of cases the test delivers a positive result when the person tested does not have the disease. When a test indicates that a particular condition is present when it is not is called a **false positive.** Let us presume that the test yields a positive result if the patient has the disease 90% of the time, and that it yields a positive result when the patient does not have the disease 10% of the time.

According to Bayes' Theorem:

$$P(Have\ disease|Positive\ Test) = \frac{P(Have\ disease)\ \times\ P(Positve\ test\ |Have\ disease)}{P(Positive\ Test)}$$

The initial probability that you have the disease is 15/100, since 15% of the people in your community have the disease. The likelihood that you get a positive test given that you have the disease is 90/100, since the test yields a positive result when the patient has the disease with 90% accuracy.

This leaves us with the task of determining the probability that you get a positive test. How are we to figure this out? Well, if you are tested, the results will be positive in two cases: if you have the disease, and if you do not. So, the probability that you get a positive test is just the probability that you either have the disease and get a positive test, or you do not have the disease and get a positive test. That is,

$$P(Positive\ test)$$

$$= P(Have\ disease\ and\ positive\ test\ or\ Do\ not\ have\ disease\ and\ positive\ test)$$

We know from earlier that,

$$P(Have\ disease\ and\ positive\ test\ or\ Do\ not\ have\ disease\ and\ positive\ test)$$

$$= P(Have\ disease\ and\ positive\ test)$$

$$+ P(Do\ not\ have\ disease\ and\ positive\ test) - P(both)$$

Since you cannot both have the disease and not have it, we can drop the last part.

Now we just need to determine the probability that you have the disease and test positive, as well as the probability that you do not have the disease and test positive. From earlier, we know that,

$$P(Have\ disease\ and\ positive\ test)$$

$$= P(Have\ disease)\ x\ P(Positive\ test\ |Have\ disease)$$

and,

$$P(Do\ not\ have\ disease\ and\ positive\ test)$$

$$= P(1 - Have\ disease)\ x\ P(Positive\ test\ |Do\ not\ have\ disease)$$

Filling in our numbers we get:

$$P(Have\ disease)\ x\ P(Positive\ test\ |\ Have\ disease) = \frac{15}{100} x \frac{90}{100}$$

and,

$$P(1 - Have\ disease)\ x\ P(Positive\ test\ |Do\ not\ have\ disease) = \left(1 - \frac{15}{100}\right) x \frac{10}{100}$$

Finally, putting together gives us:

$$P(Positive\ test) = \left(\frac{15}{100} x \frac{90}{100}\right) + \left(\left(1 - \frac{15}{100}\right) x \frac{10}{100}\right)$$

At last we can insert our values into Bayes' theorem:

$$P(Have\ disease|Positive\ Test) = \frac{P(Have\ disease)\ \times\ P(Positve\ test\ |Have\ disease)}{P(Positive\ Test)}$$

$$= \frac{\frac{15}{100} x \frac{90}{100}}{\left(\frac{15}{100} x \frac{90}{100}\right) + \left(\left(1 - \frac{15}{100}\right) x \frac{10}{100}\right)}$$

All we have left to do is calculate to find the answer.

$$\frac{\frac{15}{100} x \frac{90}{100}}{\left(\frac{15}{100} x \frac{90}{100}\right) + \left(\left(1 - \frac{15}{100}\right) x \frac{10}{100}\right)} = \frac{\frac{1{,}350}{10{,}000}}{\frac{1{,}350}{10{,}000} + \left(\left(\frac{100}{100} - \frac{15}{100}\right) x \frac{10}{100}\right)}$$

$$= \frac{\frac{1{,}350}{10{,}000}}{\frac{1{,}350}{10{,}000} + \left(\frac{45}{100} x \frac{10}{100}\right)} = \frac{\frac{1{,}350}{10{,}000}}{\frac{1{,}350}{10{,}000} + \frac{450}{10{,}000}} = \frac{\frac{1{,}350}{10{,}000}}{\frac{1{,}800}{10{,}000}} = \frac{\frac{27}{200}}{\frac{27}{200} + \frac{9}{200}} = \frac{\frac{27}{200}}{\frac{36}{200}}$$

$$= \frac{5{,}400}{7{,}200} = \frac{3}{4}$$

Given the information we supposed, the probability that you have the disease given your positive test result is only 75%.

Bayes' theorem can help us to understand what is going on in a rather infamous example in probability theory: the **Monty Hall Problem**. The Monty Hall Problem goes like this. Behind one of three doors is a prize, and you select a door. Another person, who knows which door the prize is behind, opens one of the doors you have not selected, revealing it to be empty. You may open the door you have already selected, or switch to the other remaining door. What should you do?

Many people, when confronted with this problem, answer that it does not matter whether you open the door you picked originally or switch to the other door. After all, they claim, the probability that the prize is behind each door is ½. Is this right? Let's think about the problem using Bayes' theorem.

Suppose that you pick door 1, and the other person opens door 2, revealing nothing behind it. We want to determine, given this, the probability that the prize is behind door 1, and the probability the prize is behind door 3. We can start with the probability that the prize is behind door 1.

We will let "D1 W" mean "Door 1 is the winner" and "D2 O" mean "Door 2 is opened." Plugging this into Bayes' theorem gives us:

$$P(D1\,W \mid D2\,O) = \frac{P(D1\,W)\ x\ P(D2\,O\ |D1\,W)}{P(D2\,O)}$$

Now we need to fill in the values.

"P(D1 W)"is how reasonable it is to expect the prize behind door number 1 before any door is opened. This is obviously 1/3. "P(D2 O | D1 W)" is how reasonable it is to expect door number 2 to be opened, given that door number 1 is the winner. Since either door 2 or 3 could be opened if 1 is the winner, this is ½. Finally, "P(D2 O)" is how reasonable it is to expect door number 2 to be opened, independent of any assumption about which door is the winner. Obviously, door 2 will not be opened if it is the winner, leaving 2 scenarios in which door 2 may be opened: if door 1 is the winner or if door 3 is the winner. So,

P(D2 O) = P(D1 W & D2 O) + P(D3 W & D2 O)

Obviously, before learning anything else it is reasonable to assign a probability of 1/3 each to door 1 and door 3 being the winner. If door 1 is the winner, either door 2 or 3 may be opened. If door 3 is the winner, door 2 must be opened, since the only remaining door is number 1, which you have selected.

Filling in these values gives us:

$$P(D1\,W\,|D2\,O) = \frac{\frac{1}{3} x \frac{1}{2}}{\left(\frac{1}{3} x \frac{1}{2}\right) + \left(\frac{1}{3} x\ 1\right)} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{3}} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{2}{6}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

So, given that we selected door number 1 and door 2 was opened, the probability that door 1 is the winner is only 1/3.

What is the probability that door 3 is the winner? Well, since door 2 was opened, it cannot be the winner. That means that either door 1 or 3 is the winner, which means that

P(D3 W) = 1 – P(D1 W) = 1 – 1/3 = 3/3 – 1/3 = 2/3,

Letting "D3 W" mean "Door 3 is the winner." So, given that you picked door 1 and the other person opened door 2, the probability that door 1 is the winner is 1/3 and the probability that door 3 is the winner is 2/3. You ought to switch!

However, it will be good to use Bayes' theorem again, just to be sure. So, we have

$$P(D3\,W\mid D2O) = \frac{P(D3\,W)\ x\ P(D2\,O\,|D3\,W)}{P(D2\,O)}$$

Now we need to fill in the values.

Obviously, before leaning anything else it is reasonable to assign a probability of 1/3 to door 3 being the winner. If door 3 is the winner, and you have selected door 1, door 2 must be opened. So, the probability that door 2 is opened given that door 3 is the winner is 1.

Filling in these values gives us:

$$P(D3\,W\,|D2\,O) = \frac{\frac{1}{3} x\ 1}{\left(\frac{1}{3}\ x \frac{1}{2}\right) + \left(\frac{1}{3} x\ 1\right)} = \frac{\frac{1}{3}}{\frac{1}{6} + \frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{6} + \frac{2}{6}} = \frac{\frac{1}{3}}{\frac{3}{6}} = \frac{2}{3}$$

So, if you select door number 1, and door number 2 is opened, your likelihood of winning if you stay at door 1 is 1/3, and your likelihood of winning if you switch to door 3 is 2/3. If you want the prize, it is better for you to switch.

It may be helpful to summarize this line of reasoning in a less technical way. The chance that the door you first pick has the prize is 1/3. On any occasion on which you select the winning door first, the other person could open either other door, because both are losers. So, if were to play the game over and over again, you can expect that on one out of three games you will pick the correct door the first time, and by switching you will lose. On the other hand, the chance that the door you pick first does not have the prize is 2/3. On any occasion on which you do not select the winning door first, one of the other doors must be the winner, and so the other person must open the other losing door. This means that the remaining, unopened door, must have the prize behind it. So, were you to play the game over and over again, you can reasonably expect that two out of every three games you will first pick a loser, and by switching you will win each of those times. Thus, once a door is opened, you will lose by switching 1 out of 3 times, and you will by switching 2 out of 3 times.

Work Problems 10.2

If you have a certain disease, there is a 90% chance you will break out in a purple rash. 15% of the people in your town have the disease, and 20% of people in the town have the rash. You have the rash. What is the probability that you have the disease?

Chapter 11. Inference to the Best Explanation

We are often interested in finding answers to questions of the form “Why is such and such the case?” or “Why did such and such occur?” A proposed answer to such a question is called an **explanation**. The thing to be explained, that is, the fact, state of affairs, or event about which the question is being asked, is called the **explanandum**. The proposed answer to the question, or that which is offered as doing the explaining, is called the **explanans** or the **explicans**.

Explanations are not the same as arguments. Whereas an explanation is a proposed answer to a question of the form “Why is such and such the case” or “Why did such and such occur?” an argument is a proposed answer to a question of the form “Why should I believe such and such?”

When we are interested in understanding why something is the case or why something occurred, we are not after just any explanation. Rather, we want an explanation which is correct. Just as there is no strategy that allows us to infallibly determine the truth in any given case, so there is no strategy that allows us to infallible determine the correct explanation in any case. On the other hand, just as there are often ways of determining what is most likely to be true given our evidence, there are often ways of determining which explanation is most likely to be correct. To find what explanation is most likely to be correct, we ought to consider what reasons can be offered for and against a given possible explanation. In doing this, we are engaging in a form of argument called **inference to the best explanation**. This form of argument will be the topic of this chapter.

Work Problems 11.1

For each of the following passages, identify if it is an argument or an explanation. If it is an explanation, identify the explanandum and the explanans.

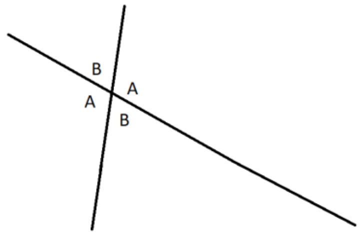
1. The car will not start because the battery is dead.
2. The battery is dead; so, the car won't start.
3. Monique aced her exam because she studies every weekend.
4. Monique must have aced her exam, since she studies every weekend.

Mathematical Explanation

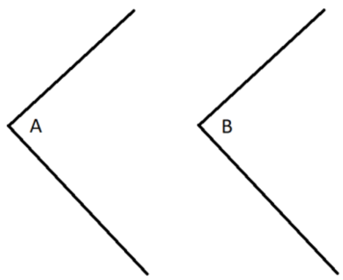
An inference to the best explanation is a form of argument with the conclusion that some possible explanation for a given explanandum is better than alternatives. Usually, the explanandum we have in mind is some event or state of affairs. It will be worthwhile to briefly discuss explanation in another area altogether: mathematics. The phrase "inference to the best explanation" does not really apply to mathematical explanations, since there need not be any one single best explanation for a given mathematical fact: multiple explanations can be just as good. This does not mean, however, that all mathematical explanations are equally good. Some explanations are better than others.

To explain why a mathematical fact is the case, one must show the fact is a consequence of other mathematical facts which we take to be true. These are called **axioms**. A mathematical explanation is good insofar as the axioms we start with are generally acceptable, and insofar as each of the steps from those axioms to the fact we want to explain are obviously valid.

As an example of a mathematical explanation, we explain why the interior angles of a triangle are the same as the angle of a line. We begin with our axioms. First, where two lines intersect, the opposite angles are equal. Thus, in the following figure, the angles marked A are equal, and the angles marked B are equal.

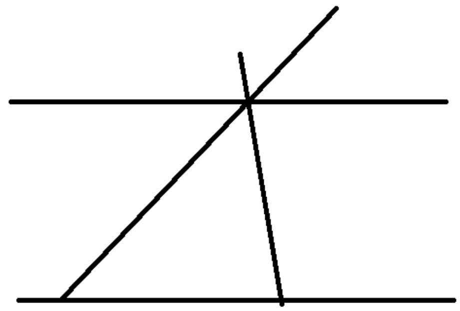


Call this axiom 1. Second, the measure of an angle does not change if you change the location of the angle. Thus, angles A and B in the following figure are equal:

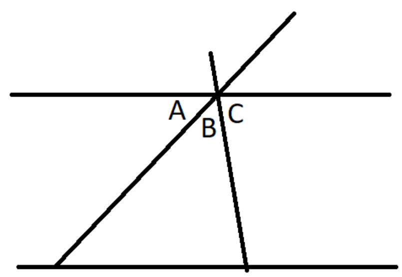


Call this axiom 2. Third, if you split an angle up into smaller angles, the sum of these angles is equal to the original angle. Call this axiom 3. Fourth, if A is equal to B and B is equal to C, then A is equal to C. Call this axiom 4.

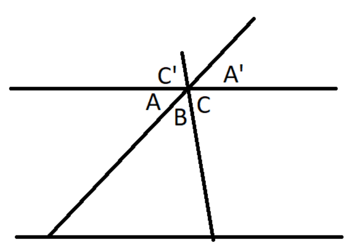
Having stated our axioms, we go on to show that the fact we want to explain is a logical consequence of these facts. First, we put a triangle between two parallel lines, so that the base of the triangle is on one line and the apex of the triangle touches the other, as such:



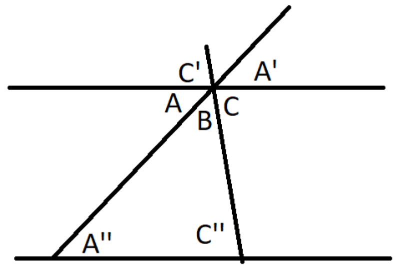
The sum of the angles marked A, B, and C in the following figure is equal to the angle of a line, by axiom 3:



In the following figure, angle A is equal to angle A’, and angle C is equal to angle C’, by axiom 1:



In the following figure, angle A’ is equal to angle A’’, and angle C’ is equal to angle C’’, by axiom 2:



The interior angles of the triangle are A'', B, and C''. The angle A'' is equal to the angle A' which is equal to the angle A, so the angle A'' is equal to the angle A by axiom 4. The angle C'' is equal to the angle C' which is equal to the angle C, so the angle C'' is equal to the angle C by axiom 4. So, the interior angles of the triangle are equal to the angles A, B, and C, which we already determined to be equal to a line. Thus, the interior angles of a triangle are the same as that of a line.

Thus, we have explained why the interior angles of a triangle are the same as the angle of a line, by showing that this is a consequence of more basic facts. Given mathematical facts which seem obviously true, our four axioms, this fact about triangles must be true.

For the remainder of this chapter, we will focus on explanations for events and states.

Prefatory Comments on Mill's Methods

Critical Thinking textbooks often have a chapter on causal arguments, where a causal argument is an argument the conclusion of which is that something A is the cause of something B. Since to identify the cause or causes of an event is to explain it, a causal argument is a kind of inference to the best explanation. Chapters on causal arguments typically contain a discussion of Mill's methods, five criteria proposed by the philosopher John Stuart Mill for identifying causes.

We could, of course, discuss how to go about identifying causes without an explicit discussion of Mill's methods. Since Mill's methods are so commonly discussed, they will be covered here too.

As is typical, the versions of Mill's methods presented here will be somewhat modified from their original versions. For one important example, instead of discussing Mill's methods as criteria for identifying causes, they will be treated instead as criteria for identifying *causal contributors*, where a causal contributor is anything which contributes causally to an event or state, whether or not it is a cause in some more precise sense. The purpose of this is to avoid having to resolve philosophical discussions about the nature of causation.

Take the following example. A match will only light in the presence of a gas like oxygen. You can strike a match in an airless container as many times as you like: it will not light. Suppose you strike a match in a normal, oxygenated environment, and it lights. What caused the match to light? Was it the striking together with the oxygen? Or was the striking of the match the cause of it lighting, and the oxygen played the role of a causal enabler? Getting clear about the nature of causation is an interesting and worthwhile philosophical project, but we want to avoid it here. Both the striking and the presence of oxygen are causal contributors to the lighting of the match.

In some discussion of Mill's methods, a distinction is drawn between necessary and sufficient causes. No such distinction will be made here. The phrases "necessary cause" and "sufficient cause" are a result of trying to understand causation in terms of material conditions. Recall from chapter 5 that in the material conditional "If A then B," the antecedent A is sufficient for the consequent B, and the consequent is necessary for the antecedent. A is sufficient for B in the sense that if A is true B must be true. B is necessary for A in the sense that for A to be true, the truth of B is required.

If A causes B, it would seem that a statement of the form “If A then B” is true, and so A must be sufficient, and so a sufficient cause, for B. If A is required for B to come about, it would seem that a statement of the form “If B then A” is true, and so A must be necessary, and so a necessary cause, for B. This reasoning is specious, however. In discussions of material conditions, the terms “necessary” and “sufficient” refer to relations between statements, the antecedent and the consequent of the consequent of the conditional. But causal relations are not between, or at least not just between statements. So, in moving from the idea that one statement is sufficient (or necessary) for another to the idea that something is a sufficient cause (or necessary cause) for an effect, we are transferring a linguistic notion to something that is not, or at least not always, part of language.

Can we make sense of the notions of “necessary cause” and “sufficient cause” on their own? The phrase “necessary cause” is strange, but we can understand it as meaning something which is required for an effect. For example, a match will not light if struck unless it is in the presence of the appropriate kind of gas. So, the presence of such a gas is required for the match to light. Thus, we might say that the presence of a gas like oxygen is a necessary cause for a match to light.

What about a sufficient cause? We can understand the notion of a sufficient cause either as a something which requires nothing else in order to bring about its effect, or as something which guarantees its effect. Let us begin with the first sense: something which requires nothing else in order to bring about its effect. Consider that striking a match does not by itself cause it to light, but only given the presence of a gas like oxygen. Reflecting, it does not seem that any cause brings about the effect that it in fact does independent of all the different background causal contributors. Thus, it would seem that a cause requires nothing else in order to bring about its effect only if we identify the cause as consisting of all the causal contributors taken together. To treat all the causal contributors together as constituting the cause of the effect, however, is to take a position on the nature of causation which we have decided to avoid.

Next, consider the second sense of a sufficient cause: something which guarantees its effect. Striking a match does not guarantee that it will light, as this effect could be prevented by removing the oxygen from the area. Because the effect which results from a causal process depends upon the various background causal contributors, a cause only guarantees its effect if we identify the cause as all the causal contributors together. This is, again, to take a position on the nature of causation which we want to avoid.

In our discussion of Mill's methods, we will sometimes talk about *causally relevant* similarities and differences between events, states of affairs, and combinations of these. It will be good to first explain what this means. In identifying causal factors, we often need to look at multiple scenarios which are very different in certain respects, or very similar in certain respects. However, any two scenarios, no matter how different, will have numerous similarities, and any two scenarios, no matter how similar, will have numerous differences. The causally relevant similarities and differences are those similarities and differences which we think could matter causally.

Compare the two extremely different scenarios of a cat on a mat and a nasty looking storm cloud forming in the distance. Both scenarios involve multiple objects, which are themselves composed of smaller objects. In both cases, those objects are physical and obey the laws of physics. Both may invoke emotional reactions in people who witness them. In both cases, a person with relevant background information about the world can make plausible predictions about the future. Both cases can be described by language. On the other hand, compare the two extremely similar scenarios of John scratching his arm on two different occasions. These at least differ with regards to the time of occurrence. Likely they differ in other regards as well: the location and intensity of the itch, what John was doing when the itch started, and what John seeing, hearing, and thinking about while scratching.

Any two scenarios will have lots of similarities and differences, but many of these are not causally relevant. In trying to identify causal contributors, you should focus on those similarities and differences which, by your best judgment, are causally relevant. Of course, it may turn out in some cases that we are wrong about what similarities and differences matter: something that we thought made no impact on what we wanted to explain might turn out to be a part of the explanans. What can we do about this? Well, if you cannot come up with an adequate explanation for an event or state using only what you think is relevant, you should start looking at those factors you did not think were relevant.

Finally, in what follows we will talk about *normal background conditions*. Normal background conditions are conditions which hold in typical cases in which the events and states we are discussing would hold, such as laws of physics holding, the atmosphere near the surface of the earth containing oxygen, and things like that. Normal background conditions are often causal contributors to an event or state; however, in explaining events we often focus on the causal contributors which do not normally hold, because those contributors are the most conspicuous.

Mill's Methods

Having completed this prefatory discussion, we turn at last to Mill's methods. In what follows, we let A, B, and C each stand for some event, state of affairs, process, or combination of these.

According to the **Method of Agreement**,

> *If multiple instances in which A is the case have no causally relevant similarities other than B and normal background conditions, either A or B is a causal contributor to the other*.

The basic idea behind the method of agreement is that if neither A nor B was a causal contributor to the other, then something else C must be. C would be some other causally relevant similarity between the instances in which A is the case. So, if there is no other causally relevant similarity than B, it must be that either A or B is a causal contributor to the other.

For example, suppose we were to find that several countries with large economies have little in common except that they have high tax rates. In that case, it would be reasonable to conclude either that a large economy causally contributes to high tax rates, or that a high tax rate causally contributes to a large economy.

According to the **Method of Difference**

> *If for two instances, one in which A is the case and one in which A is not, the only other causally relevant difference between them is that B is the case in the first case and not in the second, then either A or B is a causal contributor to the other*

The basic ideas behind the method of difference is that if neither A nor B was a causal contributor to the other, then something C would be. Were that the case, some causally relevant factor would be present where A or B was the case. Since none is, either A or B is a causal contributor to the other.

For example, suppose one country has a high literacy rate and another does not. These countries are very similar, except that in the first country a large proportion of the populace have regular access to the internet, and in the second they do not. Were this the case, it would be reasonable to conclude either that high literacy rates causally contribute to regular internet access, or that regular internet access causally contributes to high literacy rates.

According to the **Joint Method of Agreement and Disagreement**,

If multiple instances in which A is the case have no causally relevant similarities other than B and normal background conditions, and multiple instances in which A is not the case have no causally relevant similarities other than normal background conditions and the absence of B, then either A or B is a causal contributor to the other.

The join method of agreement and disagreement combines the methods of agreement and disagreement, and so its justification follows from the justification of those methods.

For example, suppose we were to find that several countries with large economies have little in common except that they have high tax rates, and suppose that several countries with small economies have little in common except for low tax rates. In that case, it would be reasonable to conclude either that a large economy causally contributes to high tax rates, or that a high tax rate causally contributes to a large economy.

According to the **Method of Residue**,

If there are two scenarios, and we have justified background beliefs that all causally relevant aspects of the first scenario except for B contribute to all causally relevant aspects of the second scenario except for A, then B causally contributes to A.

The method of residue works by process of elimination. If you can explain all causally relevant aspects of a scenario except A by some group of causes, whatever other cause that is within that group should be what causally contributes to A.

For instance, suppose that security footage shows three people leaving a bank with large, heavily stuffed bags, and later it is discovered that $100,000 is missing from the bank's vaults. Later two of the three people are captured, each with $35,000 in his bag. It is reasonable to conclude that the third person on the security footage stole the remaining $30,000.

Finally, according to the **Method of Concomitant Variation**,

If A and B always or typically vary together, one is a causal contributor to the other.

The basic idea behind the method of concomitant variation is that if A and B were not causally connected, we would not expect a change in one to regularly correspond with a change in the other. Since a change in one does regularly correspond with a change in the other, it is reasonable to conclude that the two are causally connected.

For instance, suppose that researchers discover that when the presence of a certain mineral in rivers and lakes increases, so does the presence of a certain species of amoeba. Given this, it is reasonable to conclude either that the presence of the mineral causally contributes to the presence of the amoeba, or that the presence of the amoeba causally contributes to the presence of the mineral.

One thing you may have noticed is that, with the exception of the method of residue, Mill's Methods only show that two things are causally related, but do not show which causally contributes to which. That is, they do not show the *direction of causation*. Often when using the methods we can determine the likely direction of causation by drawing from other considerations, such as our justified background beliefs. For example, suppose that all of the people who got sick at a certain restaurant had the fish entrée, but had different drinks, appetizers, and deserts. Since we know that eating spoiled food can lead to illness, but that illness does not typically cause a desire to eat fish, by the method of agreement we may conclude that they got sick because they ate the fish. In our discussion of Mill's Methods, examples were used for which the direction of causation would probably not be obvious.

But, how do we determine the direction of causation in cases in which we are not sure? Typically, we can establish causal order from temporal order: if either A or B causally contributes to the other, and A occurs before B, then it is reasonable to conclude that A causally contributes to B. Cases in which the direction of causation is not obvious are typically cases in which it is difficult to determine whether A or B came first. In such cases, we want to look for another, prior scenario in which either A or B is causally connected with something else, C.

For instance, suppose that, after having found that the presence of a certain mineral in rivers and lakes rises and falls together with the presence of a certain species of amoeba, researchers find that increases in the amount of the mineral in the rivers and lakes follows increased use of certain chemicals in nearby towns. Given that, it would be reasonable to conclude that mineral level causally contributes to the amoeba level.

Work Problems 11.2

Indicate which of Mill's Methods is being used in each of the following passages.

1. Carlos and Tom are both juniors at the same university and are taking all the same classes. Carlos studies for three hours every weekend and Tom does not. Carlos got an A on all his finals and Tom did not. So, the reason Carlos got an A on all his finals is because he studied for three hours every weekend.

2. LaToya, Manuel, and Lisa all brought cupcakes to the party. LaToya brought the vanilla cupcakes and Manuel brought the chocolate cupcakes. So, Lisa brought the strawberry cupcakes.

3. Tina found that the earlier she goes to bed, the more competent she feels while doing her job. Thus, going to bed earlier helps her feel more competent at work.

4. In a certain town, the one major difference between the small business which were successful and those which weren't is that the successful ones treated their employees well. So, either treating your employees well helps your business be successful, or having a successful business means you can treat your employees well.

5. Richard, Sarah, and Jamal are close friends. All of them are from different hometowns and had different groups of friends in high school, but all are history majors who met in class. They are friends because they met in class.

Criteria for a Good Hypothesis

Mill's methods comprise a useful set of guidelines for identifying likely causal contributors to a given explanandum. However, Mill's methods cannot always be straightforwardly applied in all cases. In our quest to find explanations, we must proceed by using our justified background beliefs about the world to form a **hypothesis** as to what the explanans for a given explanandum might be, and then test that hypothesis. If that hypothesis fails to the test, we form a new hypothesis, and test that, until we arrive at some hypothesis which is satisfactory.

A hypothesis is simply a possible explanation made for the purpose of further investigation. We form hypothesis by drawing from our justified background beliefs about the world. A reasonable person, from her experience seeking explanations for events and states of affairs, will have accumulated a catalogue of justified beliefs about how different kinds of objects, events, states, and processes can causally contribute to other events, states, and processes. Confronted with a new state or event, she or he can draw from this catalogue of beliefs to create a story of how that event or state may have come about.

Not all hypotheses are equally good. Given this, how do we evaluate a given hypothesis? Well, often we can make observations and perform experiments to test a hypothesis, which will be discussed shortly. However, there are some criteria which we can use to evaluate a hypothesis without doing this. While this seldom replaces the need for testing, such criteria can be used to narrow our possible explanations to a manageable number, and to set the stage for testing those that remain.

We want to a way to determine which possible explanans are most likely to be correct. What criteria could we use to do this? For starters, a good explanation **accounts for the details** of the explanandum. If Jill wants to understand why her car does not start, an explanas which explains the funny sound the car makes which she turns the key and the fluid leaking from the undercarriage is better than an explanans which does not account for these things. After all, unless they are simply inexplicable, something must explain the strange sound and the leaking fluid, meaning that an explanans which does not account for these is incomplete.

Second, **the candidate explanans most likely independently of the explanandum is most likely to be the actual explanans**, all else being equal. Consider how all the things which could explain the phenomenon. Suppose we were to ignore the explanandum, and ask how likely it would be for each of these candidates explanans to come about? The possible explanans which is most likely to occur, is probably the one which in fact explains the explanandum.

Third, a good explanation should **avoid positing novel entities** unless it is necessary to do so. We should prefer explanations which involve objects, processes, and events which we have independent reason to believe were likely to be present and available to produce the explanandum. If the event or process could be explained by things we already have good reason to think are present, then that is probably what explained it. For, if the explanans were to involve some novel entity, most likely we would have independent evidence of that entity, and so it would not be novel. Of course, it may sometimes happen that an explanandum cannot be explained using just entities we already have reason to think are present. In those cases, positing a new object, event, or process would be appropriate.

Finally, all else being equal, an explanation which **makes testable predictions** is better than one which does not. It would be very surprising if the explanans accounted for the explanandum but had no further effects. Starting out, the explanandum is just what we notice, and the effects of the explanans likely include more than that. Consequently, it will probably be the case that we can find evidence of these extra effects.

Work Problems 11.3

For each of the following, indicate which hypothesis offers the best explanation using the criteria discussed in the previous section.

1. Jeffrey went to sunrise coffee shop to work on his screenplay. But when he opened his backpack, Jeffrey discovered in his horror that his laptop was missing. The backpack was on Jeffrey's back for the entire bus ride to the coffee shop, but he was in a hurry packing up his stuff in his apartment this morning.

Explanation 1: A master thief stole the laptop out of Jeffrey's backpack without him noticing.

Explanation 2: Jeffrey accidentally left his laptop in his apartment.

2. Entering the kitchen after putting his daughter down for a nap, Manuel is shocked to find the pie he was letting cool on the stove has gone missing. Later, he finds a pie pan with a half-eaten pie in his son's room.

   Explanation 1: Manuel's son took the pie.

   Explanation 2: Manuel's daughter faked being asleep, took the pie, and put the half empty pan in her brother's room to frame him.

3. Richard won the election for class president despite being deeply unpopular at school and running against the extremely popular and well-liked Stephanie. Some of Richard's proposals for what he wanted to do as president were well liked by people who heard about them, but they weren't well advertised. The counting of the votes was overseen by Richard's close friend Paul.

   Explanation 1: Manuel received more votes in the election than Stephanie.

   Explanation 2: Paul intentionally miscounted the votes so that Richard would win.

4. Richard won the election for class president despite not being very popular at school. He ran against five opponents who were each incredibly popular. The counting of the votes was overseen by the current student government.

Explanation 1: Richard received more votes than any other candidate, because while the majority of votes were not for Richard, they were divided up among the other five candidates.

Explanation 2: The student government intentionally miscounted the votes so that Richard would win.

Testing Hypotheses

Once a reasonably small number of hypotheses have been identified as most likely to be true, it is often useful to test them, if possible. To test a hypothesis, we look for falsifying or confirming evidence based on its predictions.

To **falsify** a claim is to show it to be false, or likely to be false. One way of testing a hypothesis is to make a prediction about what is likely not to be the case if the hypothesis is correct, and then checking whether or not this is the case. For instance, suppose you think that the reason Jane was not in class is because she went to a party at her friend's house. If that is the case, you should expect that somebody saw Jane at the house during the party. So, if none of the people at the party remember seeing Jane, that hypothesis is likely wrong: it will have been falsified. If a hypothesis is falsified, you should reject it. By falsifying and then rejecting various available hypotheses, we can narrow the available options, with the goal of eventually reaching a single remaining explanation.

The situation is not as simple as it may seem, however. It is nearly always possible to explain away an observation which would seem to falsify the hypothesis by adding additional details. For instance, maybe no one remembers seeing Jane at the party because she came early and stayed in her friend's room, away from the other guests. In such cases, you should evaluate the modified hypothesis using the criteria discussed in the previous section. Often such modifications make the new version of the hypothesis more complicated, and will form the basis of new, testable predictions. As different versions of a hypothesis are falsified, and new modifications are made to fix it, these complications can compound, making the hypothesis less and less likely. Eventually, the cost of saving the hypothesis from being falsified becomes too high, and it should be rejected.

The other way to test a hypothesis is to look for **confirming evidence**. Confirming evidence is any observation which is likely if a given hypothesis is correct, but unlikely if it is incorrect. The mere fact that a hypothesis allows you to make an accurate prediction does not by itself provide much support for the hypothesis. It may be that what was is predicted could have come about even if the hypothesis was false. For instance, if Jane went to the party, someone probably would have seen her there an hour after class ended. But if someone remembers seeing her at the party an hour after the class, that alone would not provide much reason to think that the reason Jane was not in class was because she went to the party: she could have missed class for some other reason, and then gone to the party later. For an observation to provide strong support for a hypothesis, that observation should be unlikely if any of the alternative hypotheses are correct. For instance, if someone goes to the party immediately after class and finds Jane already there and settled, that would provide good evidence that at least part of the reason she skipped class was to go the party.

Work Problems 11.4

For each of the following scenarios, indicate one thing which would constitute falsifying evidence for the primary hypothesis, and indicate one thing which would constitute confirming evidence for the primary hypothesis as against the alternative hypothesis.

1. Stepping outside one morning, you discover your car is covered in polyethylene wrap.

    Primary Hypothesis: Your neighbor wrapped your car in polyethylene last night while you were sleeping.

    Alternative Hypothesis: A roving band of philosophers wrapped your car in polyethylene last night while you were asleep.

2. There are photographs which appear to be of astronauts on the moon.

    Primary Hypothesis: Astronauts went to the moon and took photographs of one another.

    Alternative Hypothesis: Pictures were taken of people in costumes on an elaborate set made to look like the surface of the moon.

Chapter 12. Enumerative Induction and Arguments from Analogy

In this chapter we discuss two kinds of inductive argument: enumerative induction and arguments from analogy.

Enumerative Induction

Enumerative induction is a kind of argument which contains a premise about what is the case about some observed individuals, and a conclusion about some unobserved individuals. The basic structure of an enumerative induction is like this:

N% of observed As are Bs; so, M% of As are Bs.

The observed As are the **sample**. The totality of As are the **population**. The population does not have to consist of all As: all ducks, or all people, or all apples, or whatever. It could consist of all ducks at a certain pond, all people attending a certain college, all apples in a certain barrel, and so on. B is the **target property**.

Here are some examples of enumerative inductions. First,

All the ducks we saw by the lake yesterday were mallards; so, all the ducks around the lake are mallards.

The sample is *ducks we saw by the lake yesterday*, the target property is *being a mallard*, and the population is *all the ducks around the lake*.

Second,

85% of the fish we caught from the lake were free from parasites; so, 80% of the fish in the lake are free from parasites.

The sample is *fish we caught from the lake*, the target property is *being free from parasites*, and the population is *fish in the lake*.

Enumerative induction can be usefully combined with statistical syllogism, in order to draw a conclusion about a single individual. Here is an example:

> *90% of the students I've talked to at College University enjoy sleeping in on the weekend; so, if you meet a student from College University, she or he will enjoy sleeping in on the weekend*.

This argument has two parts. The first is an enumerative induction with a missing conclusion. Here it is in full, with a plausible candidate for the conclusion included:

> *90% of the students I've talked to at College University enjoy sleeping in on the weekend; so, a significant majority of students at College University enjoy sleeping in on the weekend.*

The sample is *student's I've talked to at College University*, the target property is *enjoys sleeping in on the weekend*, and the population is *students at College University*.

The second part of the argument is a statistical syllogism, with the conclusion of the previous argument as its first premise. That argument goes like this:

> *A significant majority of students are College University enjoy sleeping in on the weekend. So, if you meet a student from College University, she or he will enjoy sleeping in on the weekend*.

The second argument is obviously inductively strong, and whole argument is strong if both parts are strong.

In this chapter we will discuss the evaluation of enumerative inductions.

Work Problems 12.1

For each of the following, identify the sample, the population, and the target property.

1. 95% of the students in the class own a computer; so, most students at the school own a computer.
2. 35% of the pets on Plain Street are fish; so, about 1/3 of the pets in the town are fish.
3. 80% of the paintings at the museum are marvelous; so, a large majority of paintings are marvelous.
4. Most college students would be benefited by becoming better critical thinkers; so, most people would be benefited by becoming better critical thinkers.
5. 80% of people surveyed enjoy Mexican food; so, most people enjoy Mexican food.

The Need for Background Information

In the introduction to this chapter, we gave three examples of arguments using enumerative induction. If we stick just with what these arguments say, there is really no good way to evaluate how strong they are. Consider the first one:

> *All the ducks we saw by the lake yesterday were mallards; so, all the ducks around the lake are mallards*.

Compare a case in which we had seen just two ducks, both mallards, and a case in which we saw twenty. It seems much more reasonable to accept the conclusion in the second case than the first.

Or take the third example:

> *90% of the students I've talked to at College University enjoy sleeping in on the weekend; so, if you meet a student from College University, she or he will enjoy sleeping in on the weekend*.

How much support the premise provides for the conclusion will depend on whether this person has talked to ten students at College University or one hundred.

What this shows is that in order to evaluate arguments using enumerative induction, we must rely on background information. This includes general background information about the world, but more specifically background information about the population, the sample, and about how the sample was obtained. In the remainder of this chapter we will discuss how such background information can be used to evaluate arguments using enumerative induction.

Representative Samples

We want to know how to evaluate enumerative inductions. That is, we want to know the conditions under which we can reliably infer from a fact about a sample to a fact about a population. One case in which we could do this was if we knew that the sample was **representative**. A sample is representative just in case

> *For any property not having to do with being part of the sample, the proportion of members of the sample which exhibit the property is approximately the same as the proportion of members of the population which exhibit the property.*

The clause, "not having to do with being part of the sample" is meant to exclude properties like "being part of the sample" or "having been observed." For, clearly, 100% of the members of the sample have been observed, but it does not follow that 100% of the members of the population have been observed.

Suppose we collect a representative sample of apples in a certain orchard, and 60% of the apples in that sample are red. Since the sample is representative, this argument will be strong:

60% of the apples observed from the orchard are red; so, 60% of the apples in the orchard are red.

After all, since the sample is representative, the proportion of apples in the sample which are red will be approximately the same as the proportion of apples in the orchard.

Unfortunately, however, this fact about representative samples does not immediate help us to evaluate enumerative inductions. For, if all we have is the sample, how do we know whether it is representative? Indeed, were we able to survey the entire population to determine whether or not the sample was representative, there would be no need to do an enumerative induction. The only evidence we have about the population is what we can tell from the sample. But how can we infer that the proportion of members of the population exhibit various properties is the same as members of the sample, unless we already know that the sample is representative?

At this juncture, background information about the population becomes crucial. For, while we may not know what proportion of the population exhibits the target property, we may well have good information about what proportions of the population have various other properties which could be related to the target property. The more of these properties our sample is representative with respect to, the more likely it is to be representative with respect to the target property.

For instance, suppose we want to know what the citizens of a certain country think about a newly proposed piece of legislation. A citizen's opinion on the proposal will plausibly be impacted by things like gender, economic status, and political and religious affiliation. It would be very strange for there to be no connection between these properties and views about the proposal, for the alternative would mean that people adopt political positions basically at random. And if people adopted political positions basically at random, it is unlikely that we would have found political positions to be associated with these other features in the past. So, if we a survey which is representative with respect to gender, economic status, and political and religious affiliation is likely to be more representative with respect to views on the proposed piece of legislation, than a survey which is only representative with respect to, say, gender.

Size and Randomness

As we saw in the last section, one way of evaluating argument involving enumerative induction is to look at whether they are representative with respect to properties which are likely to be relevant to whether individuals in the population exhibit the target property. In this section we discuss two other features of a sample that can be used to evaluate arguments involving enumerative induction: the **size** of the sample, and whether and the extent to which the sample is **random**.

A sample is random just in case

Each member of the population is just as likely to be included in the sample as any other.

If a sample is random, then as the size of the sample increases, it is likely that the proportion of members in the sample which exhibit a given property will gradually become closer and closer to the proportion of members in the population which exhibit that property. In other words, as the sample size increases, it is likely to become more representative.

To see that this is so, it may help to consider an example. Suppose that 100 waterfowl live around a certain lake: 40 geese, 30 mallards, and 20 wood ducks. We are taking a sample of these waterfowl at random, marking each one so that we do not count it twice (presume that we can do this without harming the birds). Here are the probabilities that the first bird observed is a goose, a mallard, or a wood duck:

P(goose) = 40/100 = 2/5 = .4

P(mallard) = 30/100 = 3/10 = .3

P(wood duck) = 20/100 = 1/5 = .2

Suppose the first 10 waterfowl we collect for the sample are all geese. Here are the probabilities that the next we observe is a goose, a mallard, or a wood duck:

P(goose) = 30/90 = .333

P(mallard) = 30/90 = .333

P(wood duck) = 20/90 = .222

Suppose the first 15 birds selected for the sample are all geese. Here are the probabilities that the next bird is a goose, a mallard, or a wood duck:

P(goose) = 25/85 = .294

P(mallard) = 30/85 = .353

P(wood duck) = 20/85 = .24

Here are the probabilities that the next bird is a goose, a mallard, or a wood duck if the first 20 birds observed are geese:

P(goose) = 20/80 = .25

P(mallard) = 30/80 = .375

P(wood duck) = 20/80 = .25.

What we see is, although there are more geese than mallards or wood ducks living around the lake, the more geese we include in our sample, the more and more likely it is that the next bird to be included in our sample is a mallard or wood duck.

Looking closer, we see that the next bird to be included in the sample is more likely to be a mallard than a wood duck. This makes sense, as there are more mallards than wood ducks living around the lake. However, we can also see that, were the sample to start to include lots of mallards and geese but no wood ducks, it would become increasingly likely that the next bird included in the sample would be a wood duck. So, we see that as our sample size increases it will likely become more and more representative.

Suppose at a different lake there are 75 mallards and 25 wood ducks. If we select 1 duck at random, here is the probability that it is a mallard and that it is a wood duck:

P(mallard) = 75/100 = ¾ = .75

P(wood duck) = 25/100 = ¼ = .25

If we select 2 ducks at random, the probability that both are mallards is:

P(2 mallards) = 75/100 x 74/99 = .561

The probability that both are wood ducks is:

P(2 wood ducks) = 25/100 x 24/99 = .061

The probability that 1 is a mallard and the other is a wood duck is the probability that either the first is a mallard and the second a wood duck, or that the first is a wood duck and the second is a mallard. That is:

P(1 mallard and 1 wood duck) = P(mallard then wood duck) x P(wood duck then mallard) = (75/100 x 25/99) + (25/100 x 75/99) = .379

We note that,

(75/100 x 25/99) + (25/100 x 75/99) = (75/100 x 25/99) x 2

Here, (75/100 x 25/99) is the probability of selecting one and then the other, which is multiplied by 2, since there are 2 orders in which the selection may take place.

If we select 3 ducks,

P(all mallards) = 75/100 x 74/99 x 73/98 = .418

P(2 mallards, 1 wood duck) = (75/100 x 74/99 x 25/98) x 3 = .429

P(1 mallard, 2 wood ducks) = (75/100 x 25/99 x 24/98) x 3 = .139

P(3 wood ducks) = 25/100 x 24/99 x 23/98 = .014

If we select 4 ducks:

P(4 mallards) = 75/100 x 74/99 x 73/98 x 72/97 = .31

P(3 mallards, 1 wood duck) = (75/100 x 74/99 x 73/98 x 25/97) x 4 = .431

P(2 mallards, 2 wood ducks) = (75/100 x 74/99 x 25/98 x 24/97) x 6 = .212

P(1 mallard, 3 wood ducks) = (75/100 x 25/99 x 24/98 x 23/97) x 4 = .044

P(4 wood ducks) = 25/100 x 24/99 x 23/98 x 22/97 = .004

If we select 5 ducks:

P(5 mallards) = 75/100 x 74/99 x 73/98 x 72/97 x 71/96 = .229

P(4 mallards, 1 wood duck) = (75/100 x 74/99 x 73/98 x 72/97 x 25/96) x 5 = .404

P(3 mallards, 2 wood ducks) = (75/100 x 74/99 x 73/98 x 25/97 x 24/96) x 10 = .23

P(2 mallards, 3 wood ducks) = (75/100 x 74/99 x 25/98 x 24/97 x 23/96) x 10 = .085

P(1 mallard, 4 wood ducks) = (75/100 x 25/99 x 24/98 x 23/97 x 22/96) x 5 = .013

P(5 wood ducks) = 25/100 x 24/99 x 23/98 x 22/97 x 21/96 = .001

You will notice that no particular proportion of mallards to wood ducks in any of our samples is very likely at all. This is to be expected, since there are so may combinations of ducks which might be selected. Surveying these probabilities, you will notice that as the size of our sample increases, it becomes increasingly more likely that our sample includes some combination of mallards and wood ducks, with the former comprising a larger proportion of the sample than the latter.

Moreover, consider the difference between the two largest proportions for each number of ducks selected:

One Duck: ¾ - ¼ = ½

Two Ducks: 2/2 – ½ = ½

Three Ducks: 3/3 – 2/3 = 1/3

Four Ducks: 4/4 – ¾ = ¼

Five Ducks: 4/5 – 3/5 = 1/5

We notice that the difference between them is growing increasingly smaller. The proportions which our sample is most likely to take is gradually coming to fall between a smaller and smaller range of values. Notice that the correct proportion, ¼, is between the two most probable proportions for when 5 ducks are selected for the sample: 3/5 and 4/5. What we are seeing is that, as the size of our sample increases, assuming the members of the sample are collected randomly, the sample is likely to become more and more representative.

Work Problems 12.2

For each of the following enumerative inductions, indicate if it is strong or weak, and explain why.

1. One hundred fish were examined from all different parts of the lake, at different times of the day. 15% of the fish examined were bluegill. So, around 15% of the fish in the lake are bluegill.

2. I asked ten of my friends if they liked super hero movies, and eight of them said yes. So, most people like super hero movies.

3. In a survey of 50,000 people, 1,000 per state, 90% of respondents indicated they like apple pie. The sample was representative with respect to age, sex, race, and religious affiliation.

4. In a survey of 100,000 New Yorkers, 85% of respondents indicated they prefer scrambled eggs to fried eggs. So, most Americans prefer scrambled eggs to fried eggs.

Margin of Error and Confidence Level

In previous sections, we have discussed how an argument using enumerative induction is strong if the sample is representative, and that a random sample is likely to become gradually more representative as the sample size increases. Another thing to look for is whether the study in which the sample was obtained included a **margin of error** and an associated **confidence level**. In general, the smaller the margin of error and the larger the confidence level, the more likely it is that the proportion of members of the sample exhibiting the target property is similar to the proportion of members of the population exhibiting the target property.

We can see that this is the case by considering what a margin of error and an associated confidence level represent. The margin of error and the confidence level can both be represented by a percentage. What they mean is that, if you were to repeat the survey over and over again using the same methods but different samples, the proportion of the sample exhibiting the target property would be the margin of error more or less than the proportion of the population exhibiting the target property, for the percentage of surveys equal to the confidence level.

(Note: the margin of error for a study is half the **confidence interval**. The confidence interval is the full range of values for which we can have a given confidence level, from however much below the proportion in the sample to however much above.)

For example, suppose that according to a survey with a 5% margin of error and an 80% confidence level, 60% of people enjoy Mexican food. This means that, if we were to do the survey over and over again using the same methods, but surveying different people, for about 80% of the surveys, the percentage of people surveyed who like Mexican food (assuming no one is lying) will be within 5% of the percentage of people who actually like Mexican food. It is, of course, possible that in this particular survey, the percentage of people who indicated enjoying Mexican food is much more or less than the percentage of people generally who enjoy Mexican food. But, as the confidence level is high, this is very unlikely. So, it is reasonable to conclude that the percentage of people generally who enjoy Mexican food is close to the percentage in the survey.

Arguments from Analogy

A person making an **argument from analogy**, also called as **analogical argument**, draws a conclusion about one thing from premises about similarities it has with at least one other thing. The things being compared are called **analogs**. The thing the conclusion is being drawn about is called the **primary analog**. The things to which it is being compared in the premises are called the **secondary analogs**.

Here are some simple examples of argument from analogy, with the primary and secondary analogs identified. First,

> *My friend says his Techno-co laptop is great. This laptop was made by Techno-co. It must be a great laptop as well.*

The primary analog is “this laptop” mentioned in the second sentence. The secondary analog is the friend’s laptop. Second,

> *Bob and Lisa both say their Techno-co laptop is garbage. This is a Techno-co laptop. It probably isn’t very good*.

The primary analog is “this laptop” mentioned in the second premise. The secondary analogs are Bob’s laptop and Lisa’s laptop.

Work Problems 12.3

For each of the following arguments from analogy, identify the primary analog and the secondary analog or analogs.

1. John’s car gets good gas mileage. This car is a similar model, so it must get good gas mileage.

2. The last three times Countryland entered a recession, they elected a conservative president. Countryland has entered a recession. In the next election, they will elect a conservative president.

3. Rabbits have long legs that they keep coiled up, and they can jump far. Frogs have long legs that they keep coiled up, and they can jump far. Members of this newly discovered species have long legs that they keep coiled up. So, members of this newly discovered species can jump far.

Basic Approach

A person sincerely offering an argument from analogy intends to make the case that something has a certain feature or feature, because it is similar to something else, or some things else, that has that feature. You liked this movie, so you'll like this other movie of the same genre. You ought to apply for this job, because his friend with a degree in the same field as yours got a similar job. But, why should we think that if two things are similar in some ways, that they are similar in some other way as well?

In considering this question, it is important to keep in mind that in arguments from analogy, we are comparing different things. You liked this movie, so you'll like *this other movie*. You ought to apply for this job, because *this other person* with a degree in the same field got this *different job*. Since any two things will differ in numerous ways, the primary and secondary analogs will in many ways be *dissimilar*.

Of course, the conclusion of an argument from analogy is not that certain things are similar in *all respects,* but only that they are (or are likely to be) similar in some particular respects. An argument from analogy will be strong just in case, and to the extent which, the known similarities between the analogs given in the premises give us reason to believe the analogs are also similar in the particular way specified in the conclusion: that you will like this other movie, or that you will likely get hired for this job. Thus, in evaluating an argument from analogy, we must ask whether we have good reason to believe that the similarities between the primary and secondary analogs given in the premises are likely to indicate the further similarity claimed in the conclusion. Several criteria for answering this question will be discussed in the sections to follow.

Relying on Preestablished Connections

Sometimes, a person arguing from analogy assumes that her or his audience has background beliefs about certain features being connected: that if something has this feature, it is likely to have this other feature. In such cases, the person making the argument is not actually attempting to establish a connection between the shared features of the primary and secondary analogs mentioned in the premise and the feature mentioned in the conclusion. Rather, this person is relying on her or his audience already accepting this connection.

In such cases, the effect of the argument from analogy is to draw attention to certain features of the primary analog and remind the audience that they already believe (or at least so the arguer presumes) that those features are connected to other particular features. For instance, consider the following argument from analogy:

> *Johnny's friend Mark takes piano lessons from Ronda Jones. Mark has quickly started showing real skill on the instrument, and says Mrs. Jones is a lot of fun. If Johnny is interested in learning about music, you should have him got to Mrs. Jones for lessons.*

A person making this argument would likely assume her or his audience believes the following things:

> *Friends have a significant range similar interests, likes, and dislikes*.
>
> *Good progress in a student is a sign of a good teacher*.
>
> *Often it helps a student to learn from someone he or she enjoys being around*.

If you accept these background assumptions, it is reasonable for you to accept that since Mark is becoming a good pianist from taking lessons with Ronda Jones, Johnny would also likely become a good pianist from taking lessons with her, assuming he is interested.

Work Problems 12.4

For each of the following arguments from analogy, identify some things a person making the argument likely presumes her or his audience believes.

1. John's car gets good gas mileage. This car is a similar model. In fact, the two models only differ in minor aesthetic ways. So, this other car must get good gas mileage.

2. The last three times Countryland entered a recession, they elected a conservative president. Countryland has entered a recession. In the next election, they will elect a conservative president.

3. Rabbits have long legs that they keep coiled up, and they can jump far. Frogs have long legs that they keep coiled up, and they can jump far. Members of this newly discovered species have long legs that they keep coiled up. So, members of this newly discovered species can jump far.

Establishing Connections: The Basic Challenge

In the previous section we discussed cases in which an argument from analogy relies on the audience's background beliefs about certain features being connected. What about arguments from analogy which are not like this? We have to consider arguments from analogy where the person making the argument cannot rely on her or his audience already believing if the something has the features shared by the primary and secondary analogs in the premises, it is also likely to have the feature ascribed to the primary analog in the conclusion.

In such cases, the connection between these features must be established within the argument itself. How is this to be done? Assuming the premises of the argument are true, each secondary analog will share at least one feature with the primary analog, and also exhibit the feature which we ascribe to the primary analog in the conclusion. For each secondary analog, either it just so happens that it exhibits both features, or there is a connection between these features, in the sense that if something exhibits the one, it is more likely to exhibit the other. An argument from analogy of this type will be strong just in case, and the extent to which, the information provided in the argument, together perhaps with some justified background beliefs, provides us with good reason to reject the second option.

Establishing Connections: An Example

When evaluating such arguments from analogy, what should we look for? To answer this question, it may be worthwhile to imagine a case in which we must reason analogically but cannot rely upon our background beliefs about what features are connected to one another. Suppose that scientists have discovered extraterrestrial plant life, and a large number of samples have been brought to earth (ethical questions could be raised, but we will put them aside). Scientists have been able to determine that several of the species are poisonous, and several are nonpoisonous. Unfortunately, funding for the program was slashed before all species could be classified. You have been tasked with trying to determine, as best you can, which of the remaining plants are poisonous and which are nonpoisonous, but you cannot perform any sort of chemical analysis on the samples.

How would you proceed? Well, you could eat some of the plants and see what happens, but maybe we should leave that as a strategy of last resort. Consider what information you have available: you know that certain plants have already been determined to be poisonous, and others nonpoisonous. Presume that you have access to a high-quality specimen of each species and a database indicating whether it is poisonous, nonpoisonous, or unclassified. How might you use this information to come to a reasonable conclusion about which of the unclassified plants are nonpoisonous and which are poisonous?

If there are any features which are shared roughly equally between poisonous and nonpoisonous plants, then clearly the fact that an unclassified plant has those features does not help you to determine whether it is nonpoisonous or poisonous. On the other hand, suppose that there is some feature which appears in several poisonous plants, but which is absent from all plants which we have determined to be nonpoisonous. That would give us some reason to think that the feature is an indicator of a plant's being poisonous. After all, if this feature was not connected with a plant's being poisonous, then as we examine more and more plants, we would reasonably expect to find non-poisonous plants which exhibited this feature as well. So, if an uncategorized plant has this feature, this fact would constitute some reason to think the plant is poisonous.

Continuing, it may be that there is some feature which appears frequently in poisonous plants and only rarely in nonpoisonous plants. As the number of poisonous and nonpoisonous plants examined increases, the less likely it would be to just so happen that a very great proportion of poisonous plants have this feature and a very small proportion of nonpoisonous plants. This would give us some reason to think the feature is typically, though not always, connected with a plant being poisonous. So, if an uncategorized plant has such a feature, this would constitute some reason to think the plant is poisonous.

Establishing Connections: Evaluating Arguments from Analogy

In the previous section, we discussed a hypothetical case in which you must rely on analogical reasoning without the aid of background beliefs about what features are connected to one another. We can use this example as a guide for how to evaluate arguments from analogy. Although these criteria will apply to analogical arguments generally, we are most specifically interested in arguments in which the arguer is trying to establish a connection between the features shared by the primary and secondary analogs in the premises and the feature assigned to the primary analog in the conclusion within the argument itself.

In a good argument from analogy, we will be given reason to believe that the features shared by the primary and secondary analogs in the conclusion are connected with the feature assigned to the primary analog in the conclusion. From our discussion in the previous section, when evaluating arguments from analogy we can consider the following:

The number of Secondary analogs -

The greater the number of secondary analogs, the stronger the analogy. The more things like the primary analog which also exhibit the feature ascribed to the primary analog in the conclusion, the feature which we are ultimately interested, the less likely it is merely a coincidence that they have this feature in common.

The number of interesting features shared between the primary and secondary analogs –

The more interesting features shared between the primary and secondary analogs, the stronger the argument. By an interesting feature is meant a feature which is in some way peculiar. In a strong argument from analogy, the features shared by the premises and the conclusion will solely or predominantly appear in things with the feature ascribed to the primary analog in the conclusion. Common features are likely to be exhibited both by things which exhibit the latter property and things which do not.

The number of interesting differences between the primary and secondary analogs –

The more interesting features which secondary analogs have and primary analogs lack, or secondary analogs lack and primary analogs have, the weaker the argument. Because these are interesting features, in the sense explained previously, they may be connected to the secondary analogs exhibiting the feature we are ultimately interested in: the feature assigned to the primary analog in the conclusion. Thus, differences of this kind between the primary and secondary analogs weakens the analogy.

The amount of diversity among the secondary analogs –

The greater the secondary analogs differ among themselves with regard to apparently uninteresting features, the stronger the analogy. The more diverse the secondary analogs are, the less likely it is for features we think to be uninteresting, in the sense explained previously, to be connected with the feature we are ultimately interested in. This provides support for believing the interesting features shared by the primary and secondary analogs in the premises really do serve as reliable indicators for the presence of the feature ascribed to the primary analog in the conclusion.

In evaluating arguments from analogy, it is best to consider these criteria in combination. An argument from analogy with few interesting similarities between the primary and secondary analogs will be weak, even if the number of secondary analogs is large, for instance.

Work Problems 12.5

Evaluate the following arguments from analogy.

1. Bob, Ralph, Rachel, Chelsea, Carlos, and Sam, while all being from different hometowns and having different part-time jobs, are all recent friends of Stephanie who go College State University. Marcus is also a recent friend of Stephanie's. So, Marcus attends College State University.

2. Cars, trucks, and jeeps of different models with the new turbox engine all get great gas mileage. This other car has the new turbox engine. So, this other car will get great gas mileage.

3. Plant A has white flowers, thorns, and in late spring small, purple berries. While plant B has no thorns, it does have white flowers. Plant B will grow small, purple berries in late spring.

Chapter 13. Arguments from Authority

A great deal of what a typical person believes are not things about which she or he has directly examined the evidence for her or himself. As a result, a significant portion of what a person believes are things which she or he has heard or read at some point or other. That is, you believe many things because other people say that they are true.

This makes sense. If you were to form beliefs only on those topics about which you could directly examine the evidence for yourself, you could only believe things about which you have direct sensory experience, and things you can figure out by thinking carefully, such as mathematical facts. By this standard, most people would be forbidden from holding beliefs about all sorts of things: how old the earth is, the likely impact of a certain policy proposal, what strains of flu virus people should be inoculated against this year, what is going on in other parts of the world, and so on. To restrict your beliefs in this way would greatly deprive you of the ability to make many kinds of informed decisions.

An argument in which the reason offered to believe the conclusion is that some person or group endorses the conclusion is called an **argument from authority**. Arguments of this kind will be the topic of this chapter.

The Problem, and Classifying Cases

In attempting to give an account of how to evaluate arguments from authority, we are confronted with a dilemma. For much of what we believe, we cannot verify that it is correct for ourselves, but must depend on what others tell us. On the other hand, if you cannot verify what is being claimed for yourself, how can you determine whether or not the claim is reliable?

To answer this question, we need to think both about how a claim could fail to be reliable, and in why and to what extent in a given case you are unable to verify the claim for yourself.

A claim could fail to be reliable either if the person making the claim is being dishonest, or if the person is reporting her or his genuine beliefs on the matter but is mistaken. So, in evaluating arguments for authority, we want to look for ways to assess these two possibilities.

Next, we can discuss different reasons you may be unable to verify a claim for yourself. This will depend on the kind of claim being made. For one, there are **claims about how things seem** to that person. It would seem that a person cannot be mistaken with respect to claims of this sort: if you seem to see a mountain lion in the field, then you do, indeed, seem to see a mountain lion in the field. It is possible, however, to lie about how things seem.

Why are you unable to verify claims about how things seem to another person? The answer is, because you cannot have someone else's experiences, but only your own. Except, this does not show that you cannot verify a claim about how things seem to someone else. Rather, it shows that you cannot do so definitively. It may be reasonable, in some cases, to think that a person is lying about how things seem to him. If you thought someone might be lying about how things seemed to him, you might draw from your own experience to imagine how things might seem to someone in that person's situation. That being so, it does not seem that people typically lie about how things seem to them.

Next, there are **claims about things experienced first-hand** by that person, but for which you were not present for. Such claims are distinguished from those of the previous class, since these are claims about what is the case, rather than how things seem to someone. The claim that you seem to see a mountain lion in the field is a statement of the previous kind, whereas the claim that you in fact see a mountain lion in the field is a statement of this second category.

When making claims of this kind it is possible to lie, as well as to be honest but mistaken. It may seem to you that there is a mountain lion in the field, but you might just be misjudging the size of a housecat in the distance. If that is the case, then you will be mistaken when you claim that you do, in fact, see a mountain lion. What seems like a face in the shadows might just be a tree. In that case, you will be mistaken when you claim that you in fact see a face in the shadows.

The obstacle to your verifying such claims is that you were not present to witness what happened for yourself. While this is indeed the case, there may still be things you can do to assess the claim. Perhaps you could consult someone else who was present? If that is not possible, you may have been in a similar situation yourself, and if so, you could use your own experience as a basis for evaluating the claim.

Next, there are **third-party reports**, or reports about how things seemed to some other person, or what some other person experienced first-hand. That is, claims about what someone else heard or saw, or what happened to them. Many news reports are like this. Someone witnesses an event. This person is interviewed by a reporter who then writes a story, so what appears in the news is what the reporter claims the witness saw. Or, it may be that a different reporter for a different news outlet writes a piece based on the first report. Indeed, there may be a great number of steps between the person who witnessed the event and the person telling you what that first person heard or saw. No matter the number of steps, any person passing along the information could lie about what the first person said. More, she or he could be mistaken, as a result of misunderstanding what she or he was told.

In this kind of situation, you might be able to converse with the original witness for yourself, though this could require a large expenditure of time and resources. You would have to identify who the person was, find out how to get into contact with her or him, find a way to address any linguistic barriers between the two of you, and so on. All this could be very costly. Even worse, the original witness might no longer be alive. But even in this case, there might be different accounts of the original claim which you could compare.

Finally, there are **reports about the meaning or importance of something** someone witnessed. For instance, what does this study or experiment show? Within this category are **theoretical statements**, or statements involving theories, such as in science. A person could be mistaken about the significance of a study or experiment, because she or he does not know how to properly interpret the data, or because she commits a mistake in reasoning. Furthermore, a person who knew how to correctly interpret the data (or who thought this), could also lie about what that data reveals.

The reason it is difficult to assess such claims for yourself is that you may not have the background knowledge and training to appropriately interpret the data yourself, or to adequately understand the theory. In principle you could learn enough about the relevant field to acquire this ability. Becoming an expert in many fields takes years of intense study, though, which is often not feasible.

Work Problems 13.1

For each of the following passages, indicate whether it is a claim about how things seem, a claim about something experienced first-hand, a third-party report, or a report about the meaning or importance of something. Then indicate whether the claim could be unreliable because the person making it is lying, mistaken, or both.

1. John says a certain study reveals that people with blue eyes have less than average intelligence.
2. Jane says she saw Tom run a 4-minute mile.
3. Tom says Rachel said she had a date with the bass player of her favorite band.
4. Josephine says she thought she saw an alien spacecraft when she was going for a walk in the countryside.

General Considerations for Evaluating Arguments from Authority

In the previous section, we noted that there are basically two ways a person's claim could be unreliable: if that person is lying, or if that person is honest but mistaken. We also saw that there are many different kinds of claims for which we rely on the authority of others, and that in many of these cases there are ways to assess these claims, but that doing so is often very difficult. In this section we will discuss evaluating such claims in more detail.

As has been discussed, one of the major reasons we rely upon arguments from authority is that finding and assessing the evidence for ourselves can be extremely costly in terms of time and resources. This is also an important factor in why it can be difficult to evaluate arguments from authority. Because of this, when considering arguments from authority, it is important to begin by evaluating the importance of the claims being made. What are the costs of believing this claim if it is false? What are the costs of failing to believe this claim if it is true? The higher the costs involved, the more effort you should put in to trying to assess the claim. If the costs are negligible, you might just refrain from taking a position at all.

In evaluating an argument from authority, your goal is to seek out reasons for and against the reliability of the person (or persons) making the claim. In general, if the considerations in favor of the person being reliable about the claim are much stronger than the considerations against, it is reasonable to accept the claim. If the considerations against the person being reliable about the claim are much stronger than the considerations in favor, it is reasonable to be suspicious of the claim, and in some cases even to reject it. The following are several things to consider.

Are the source's claims about the topic consistent? Inconsistent claims, that is, claims which directly contradict one another, cannot all be true. So, if a person says inconsistent things about some topic, at least some of what she or he is saying must be false. Furthermore, the fact that the person is saying false things about the topic gives you some reason to suspect that in making these claims, she or is interested in something other than just the truth. That being said, the fact that a person makes inconsistent claims at different times does not necessarily mean she or he is being intentionally deceptive. John might say that Sally is angry with him one moment, and later say she was not angry with him, but just in a bad mood generally from being hungry and tired. These are inconsistent claims, and they cannot be true. However, it seems likely that John changed his mind about what happened rather than lied.

Has the source been reliable about the topic in the past, and attempted to address errors? Everyone gets things wrong now and then. A person who is primarily interested in finding the truth is willing to admit her or his errors and work to try to prevent the similar mistakes from happening in the future. A person with a record of being wrong who is hesitant about owning up to these errors, and who makes little or no effort to address them, is likely interested in something other than the truth. Just as well, if a person does admit her or his errors and tries to address them but still has a history of getting things wrong, you have reason to question that person's reliability with regards to that topic.

Is there corroborating or disconfirming evidence for the claim? Is there evidence to back up what the person says, or evidence to challenge the claim? Seeking out such evidence may be costly, but if the costs of being mistaken about the claim are high enough it might be worth it.

Is the source primarily motivated by personal benefit? People rarely lie arbitrarily, but because it benefits them in some way. Sometimes people lie in an attempt to get out of trouble. Sometimes people lie for the pleasure of having fooled someone. This is not necessarily malicious: people often exaggerate to their friends for fun, where no harm is intended. Sometimes people lie in order to manipulate others. The mere fact that it would benefit someone for you to believe her or his claim, however, does not necessarily mean that the person is lying. After all, it is often beneficial for someone if others believe what is true. If your believing a claim would likely benefit the person making it, do not automatically reject the claim, but consider what that person's primary motivation seems to be.

Does the source engage in rhetorically questionable tactics? Back in chapter 1, we discussed how sometimes a speaker might try to stir up peoples' emotions, and appeal to their desires to be a part of some groups, or to not be part of some groups, in order to get them to accept her or his claims without thinking critically about them. We will discuss more tactics like this in chapter 15. A person who engages in behavior which has the intended effect of limiting good critical thinking is probably interested in something other than the truth.

Is the claim endorsed or rejected by independent sources? Do others who have no particular incentive to agree with the person also endorse the claim? Do others who have no particular incentive to challenge the person reject the claim? If multiple people who independently examine the same topic come to the same conclusion, it is at least somewhat likely that the conclusion is close to the truth. People who lie independently are likely to lie in different ways. People who make mistakes independently are likely to make different mistakes. In asking this question, however, it is important to consider factors that could be motivating multiple people who are otherwise unconnected. Do people accept this claim simply because it is generally believed, or do multiple people accept it because they have independently investigated it?

Work Problems 13.2

1. Give an example of a claim which it is important to have correct beliefs about.
2. Give an example of a claim for which there will likely be little cost to refraining from taking a position on.
3. Give an example of something that might incentivize a person to lie about an important claim.

Experts

Earlier we classified different kinds of claims for which you might rely on another person's authority. This included **reports about the meaning or importance of something**, including **theoretical statements** and statements about the significance of a study or experiment. Arguments from authority involving such claims deserve special attention. The reason for this is that often correctly interpreting the meaning of a study or experiment, or understanding a theory, requires specialized knowledge about the topic. A person with specialized knowledge about a topic is an **expert** on that topic. In this section we will discuss things to consider when evaluating such arguments.

The first thing to consider is, how to you determine that someone is an expert in a given field? One important sign that a person is likely to be an expert is if she or he **has an advanced degree in the relevant field**. In order to earn an advanced degree in a given field, a person must spend years studying that field, with the guidance and criticism of other experts in the field.

Another sign that someone is likely to be an expert on a topic is if she or he **has published an article on the topic in a peer reviewed journal**. In a **peer reviewed journal**, articles are only accepted for publication after being reviewed and accepted by experts in the field. To be published in a peer reviewed journal therefore means that experts in the field believe you are contributing something important. This is difficult to do for someone who has only a casual understanding of the topic. Best is when the journal uses **double-blind review**, meaning that the author of the article does not know who reviews it, and the reviewer does not know who wrote it. This helps to eliminate favoritism in which articles are selected for publication. It is possible for bad articles to be published in good journals, but for high quality journals this is the exception.

Yet another sign that someone is an expert on a topic is if that person is **aware of other experts, their points of view, and points of controversy**. To be an expert in a field requires having read and thought about what other experts in the field have said, whereas a nonexpert has little incentive to do so. A person who is not familiar with these things likely does not have special insight into the topic. Especially with complicated topics, it is good not to give too much credence to the claims of people who are only casually familiar with the material.

Once you have concluded that someone is an expert in a given field, there are further things to consider before accepting what that person says. First, **is the claim being made within that person's area of expertise?** Just because someone is an expert in one field, does not mean that person has special insight into topics outside of that field. Before accepting a someone's claims because that person is an expert, make sure she or he is an expert on the right topic.

Second, **is there consensus among the relevant experts with respect to the claim**? Sometimes there may be a great amount of disagreement on certain issues between experts in the same field. This likely indicates that it is not clear which view the available evidence best supports. In such cases the fact that an expert endorses a claim, even in expert in the relevant field, does not give you much reason by itself to think the claim is true. On the other hand, the less controversial the claim is among the relevant experts, the more reasonable it is for you to accept it.

Work Problems 13.3

1. Steve says early humans ate nothing but grass. You should believe him, because he has a PhD in philosophy.
2. A group of renowned economists say the new tax bill would weaken the economy. So, if we don't want to weaken the economy, we should reject the new tax bill.
3. Bob Smith, a successful businessman, says the new tax bill would strengthen the economy. If we want to strengthen the economy, we should support the new tax bill.
4. Marcus came up with an idea for a plane. Jill Jeffries, who has advances degrees in engineering and aeronautics, says it would never get off the ground. Marcus's plane wouldn't work.

5. Monique Alexander, a professor of physics at College University, says that any serious physicist will say the chance of the new particle accelerator creating a blackhole and destroying the earth is less than the chance of being struck by 20 bolts of lightning. So, we don't have to worry about the new particle accelerator creating a blackhole and destroying the earth.

Self-Criticizing Investigative Systems

We will conclude this chapter by discussing one more thing to consider when evaluating an argument from authority: whether or not the claim was produced within something that might be called a self-criticizing investigative system. A self-criticizing investigative system is a system the members of which seek out the truth on some range of topics, in significant part by constructively challenging the efforts of others in the system to do the same.

A good example of a self-criticizing investigative system is science. Science proceeds according to what is sometimes called the **scientific method**. Although the details very from field to field, the basic outline of the method goes like this:

1. Some phenomenon is observed.
2. A hypothesis is formed to explain the phenomenon.
3. Lots of people look for evidence for and against the hypothesis, either by collecting data or by performing experiments.
4. The hypothesis is modified in light of the evidence, or it is rejected and replaced with a new hypothesis.
5. Lots of people look for evidence for and against the modified hypothesis or the new hypothesis.

6. This process is repeated over and over again, until we arrive at a hypothesis which is consistently supported by the relevant available data.

In assessing hypotheses, scientists rely on many of the patterns of reasoning that have been discussed in this book.

Science proceeds in part by a process of elimination. In fact, sometimes this is used as a basis for a challenge to science: scientists have been wrong about a lot of things in the past, why should we believe what scientists say now? In a way this objection is correct: many hypotheses which scientists accept today are probably not exactly right. But this objection to science is based on a misunderstanding of how science works. We know scientists have been wrong in the past, because the scientific method has revealed certain hypotheses those scientists accepted to be very likely wrong. The fact that scientists have gotten things wrong counts in favor of the scientific method, since it is by using the scientific method that we learned those scientists were mistaken. And as more and more mistaken hypotheses are eliminated, the more reasonably confident we can be thinking that the remaining hypothesis are closer to the truth.

It is important to note that, although an individual scientist may do research, form hypotheses, and design and perform experiments by her or himself, science is very much a collective enterprise. Whenever one scientist conducts a study or performs an experiment, other scientists try to replicate the study or the experiment and see if they get the same kind of results. If the results cannot be replicated, scientists start to question the original study or experiment. The fact that science works this way, with different people testing one another’s methods and conclusions, is a crucial factor in why science functions so well in expanding our understanding of the world. Science works in a way which eliminates hypotheses which are not well supported.

To see this, consider an example. Seeking scientific glory, Jake makes up a study, does enough research to learn what a good discussion of a study in that field looks like, writes an article on the made-up study, and submits it to a scientific journal. Suppose the journal publishes the fictional study. Were this to be discovered, would the reputation of the journal be in serious jeopardy? Maybe, but it probably should not be. If the description of a study is well-written, it is not clear how someone is supposed to determine whether it was fictional or not, just from the description itself. Within the description of the study will be the methodology used, the data found, and possibly an interpretation of the data. The methodology described could be good, and the interpretation reasonable, even if the data is completely made up.

Does this mean that Jake's fake study becomes accepted science? Thankfully, it does not. For after the study is published, other scientists will conduct similar studies to see if they get the same results. Others will perform variations on the study, to examine related questions. Scientists will begin to question Jake's study if they cannot successfully replicate the results. While seeking scientific fame, Jake instead gains scientific infamy. Serious scientists will proceed to ignore Jake in the future.

The scientific method functions to filter out bad hypotheses. A bad hypothesis does not need to be born from intentional deception, such as with a fake study. People can simply make mistakes. Individual scientists are people, and people are flawed creatures with different kinds of biases. An individual scientist may fail to see certain patterns in the data or may fail to think of certain interpretations of the patterns she or he does see. A scientist may fail to think of a possible experiment to test a hypothesis, or the scientist may fail to fully understand the significance of the results of an experiment or study. These lapses will be accounted for by the efforts of other scientists, collecting their own data, performing their own experiments, and interpreting the results for themselves. This is not because these other scientists are infallible, but rather because different scientists are likely to make different mistakes, and a mistake made by one scientist is likely to be corrected by another. As time passes, we get closer and closer to what the data actually indicates.

By the way, this is why it is important to have people from different backgrounds doing science: it is often very difficult to know in what ways and to what extent the way you think is influenced by facts about you and your social environment. Although one should try to put personal biases aside when seeking to learn about the world, to do this entirely is an extremely difficult, if not futile endeavor. By having people from different backgrounds participate, the scientific method functions to account for those biases.

In our earlier example, the results of Jake's fake study are eventually dismissed because other people, real scientists, are participating in the scientific method: they are engaging in and contributing to a self-criticizing investigative system. What if they were not? What if all, or even a vast majority, of so-called "scientists" were frequently faking the results of studies and experiments. In that case, science would fail to operate as a self-criticizing investigative system. As a result, we would have no special reason to accept what scientists say.

On reflection, the idea that this is happening in the real world is wildly implausible. People are taught to do science by learning about the scientific method. Students are made to read and think about the results of studies and experiments performed by other scientists and tasked with conducting their own studies and performing their own experiments. Students continue in their studies in part because they believe in the method. Now imagine that, after rigorous scientific training over many years, the student is invited to a gathering. The student is taken into a room, and the doors are shut. Someone rises and reveals the truth: it is all a farce, we don't really use the scientific method at all, but just make up the data as we like! How would the student, who had spent so much time doing studies and performing experiments, react to this pronouncement? If "science" was just a fake, there would surely be a huge outcry from the upcoming scientists of the new generation.

Alternatively, it could be that a significant portion of scientists are faking the results of studies and experiments, but they are doing so independently, and not as part of an organized effort. Again, this seems exceedingly unlikely to be the case in the actual world. The student of science undergoes years of training during which the importance of collecting data, performing experiments, and looking for falsifying and corroborating evidence is emphasized again and again. We might reasonably expect that some scientists will deviate from the scientific method and engage in questionable behavior, such as faking data. But it seems preposterous to think a large portion would. And so long as the predominant number of scientists follow the scientific method, science will function as a self-criticizing investigative system, and so will work to eliminate mistakes and fakes.

This reveals another way in which science is a self-criticizing investigative system. Learning about science includes learning about the scientific method, which incentivizes people who genuinely believe in the method to continue in the field and for people who do not to leave it. Fools and frauds will nevertheless sometimes become respected scientists. But the nature of scientific practice itself wards against this happening with excessive frequency.

What is the significance of all this for evaluating arguments from authority? For one, if a great many scientists who specialize in the same field all agree about something in that field, that is a good reason to think that what they say is likely to be true, or close to true. And this holds for other self-criticizing investigative systems, as an ideal journalism team might be. More generally and perhaps more importantly, it means that while you cannot investigate everything for yourself, it is still reasonable for you to believe some of the claims made by people operating within self-criticizing investigative systems.

An important further consequence of this, is that even if you accept a claim on authority without checking to verify it for yourself, it is important that someone does so. If you are interested in knowing what is most likely to be true given the available evidence, then you should want the claims of authority figures to be tested and evaluated. While none of us can do this all of the time for all claims, almost all of us can do this some of the time for some claims. By doing this, we help foster an environment which favors claims which are best supported by the evidence.

Part IV. Fallacies and Rhetorical Trickeries

Chapter 14. Fallacies

The purpose of this book is to give you the tools to practice more effective critical thinking. So far, this has largely consisted in identifying different kinds of arguments and discussing techniques for evaluating arguments of those kinds. In this part of the book, we will discuss some ways that a person can reason badly, and ways that a person's genuine attempts to reason well can be sabotaged. By learning about these things, you guard yourself against them, and to help others be better critical thinkers as well.

The topic of this chapter will be **fallacies**. Broadly speaking, a fallacy is any argument the premises of which fail to provide good reason to believe the conclusion. In discussing fallacies, we are often more interested in fallacies in a narrower sense: an argument the premises of which fail to provide good reason to believe the conclusion, but which is often psychologically persuasive. An argument which commits a fallacy is said to be **fallacious**.

Formal Fallacies

The topic of Part II of this book was deductive arguments. In those chapters we made a distinction between the **form** and the **content** of an argument. The form of an argument is the basic structure of the argument, independent of what the argument is about. This latter part, what the argument is about, is its content. We can generate an argument from an argument form by taking that form and inserting some content. An argument form is valid just in case, no matter what content we insert into it, the resulting argument will be valid. An argument form fails to be valid just in case there is some content we could insert into that form, such that the resulting argument is invalid. Any argument that does not have a valid form is, in the broad sense of fallacy, a **formal fallacy**.

As there are an infinite number of arguments forms which are not valid, so there are an infinite number of formal fallacies, in the broad sense of an argument in which the premises fail to provide a good reason to believe the conclusion. Some such argument forms are used fairly frequently, however, and people are often convinced the conclusions of arguments with those forms. We will discuss two examples here.

The first formal fallacy we will discuss is **affirming the consequent**. An argument commits the fallacy of affirming the consequent just in case it has the following form:

If P then Q	(Premise)
Q	(Premise)
P	(Conclusion)

Recall that there are special names for the statements comprising a material conditional. When a conditional is written as an "if… then" statement, the part after the "if" is the antecedent, and the part after the "then" is the consequent. In a statement of the form "If P then Q," the antecedent is "P" and the consequent is "true". An argument affirms the consequent just in case one of its premises is a material conditional, and the other premise is the antecedent of that conditional statement. Thus, the antecedent of the conditional is affirmed in the other premise.

Here is an example of an argument which affirms the antecedent:

If John's pet is a dog, then John's pet is a mammal.

John's pet is a mammal.

So, John's pet is a dog.

With a little reflection, it is fairly obvious that this argument is not valid. Remember, an argument is valid just if it cannot both be the case that the premises are all true and the conclusion is false. There is a way for both of these premises to be true and the conclusion false. Suppose John does not have a pet dog, but instead has a pet cat. The first premise is true: if John's pet is a dog, it is a mammal. John's pet is not a dog, but the premise is still true. The second premise is also true: John's pet is a mammal, since cats are mammals. The conclusion is false, since John does not have a pet dog.

The reason people often accept the conclusion of arguments which affirm the consequent, is that arguments with this form can be easily confused with arguments with the form **Modus ponens**. Modus ponens was introduced in chapter 7 as one of the rules of inference for propositional logic. Modus ponens refers to the following argument form:

If P then Q	(Premise)
Q	(Premise)
P	(Conclusion)

Any argument with the form modus ponens is valid.

An example of an argument with the form Modus Ponens is

If John's pet is a dog, then John's pet is a mammal.

John's pet is a dog.

So, John's pet is a mammal.

Take a moment to compare the argument form modus ponens with the form affirming the consequent, so that you recognize the difference.

The second formal fallacy we will discuss is called **denying the antecedent**. An argument commits the fallacy of denying the antecedent just in case it has the following form:

If P then Q (Premise)

Not P (Premise)

Not Q (Conclusion).

An argument denies the antecedent just in case of its premises is a material conditional, and the other premise is the negation of the antecedent of that conditional. Thus, the antecedent is denied in the other premise.

Here is an example of an argument which denies the antecedent:

If John's pet is a dog, then John's pet is a mammal.

John's pet is not a dog.

So, John's pet is not a mammal.

A little reflection reveals this argument is obviously invalid. Suppose John's pet is a cat. The first premise is true, since if John's pet is a dog, it is a mammal. The second premise is also true, as John's pet is not a dog. But, the conclusion is false, since John's pet is a cat and cats are mammals.

The reason people often accept the conclusion of arguments which affirm the consequent, is that arguments with this form can be easily confused with arguments with the form **modus tollens**. Modus tollens was introduced in chapter 7 as one of the rules of inference for propositional logic, and refers to the following argument form:

If P then Q (Premise)

Not Q (Premise)

Not P (Conclusion)

Any argument with the form modus tollens is valid.

An example of an argument with the form modus tollens is as follows:

If John's pet is a dog, then John's pet is a mammal.

John's pet is not a mammal.

So, John's pet is not a dog.

Take a moment to compare the argument form modus tollens with the form denying the antecedent, so that you recognize the difference.

Work Problems 14.1

Informal Fallacies

The topic of this section if **informal fallacies**. An argument is an example of an informal fallacy just in case it is fallacious for some reason other than its logical form. Several such fallacies are so well known that they have been given names. Some of these will be discussed in the sections to follow.

Informal Fallacies Part I

The **genetic fallacy** is arguing that something should be accepted or rejected in virtue of features about its origin which are independent from whatever good or bad qualities it may have now. Suppose that in attempting to poison his rival, John inadvertently developed an excellent headache remedy. You would be committing the genetic fallacy where you to argue that the remedy should not be used, since it was originally intended as a poison.

It is important that the reason offered for accepting or rejecting something is irrelevant to its good and bad qualities now. To argue for the conclusion that you should not eat a certain food because there was a contamination at the packaging plant would not be an instance of the genetic fallacy. After all, if it is true that some of the food was contaminated, that would be a very good reason not to eat it.

Determining whether an argument of this kind is fallacious or not can be tricky, and you may have to consult background information. Suppose someone argues that, while Senator Smith touts the benefits of his new policy proposal, you should not believe him, since he is a member of the untrustworthy Purple Party. Does this person simply disagree with the policies of the purple party? Then the argument is fallacious. On the other hand, if there is an established history of purple party members making false claims about the effects of their policies, that could give you reason to be skeptical of the positive claims about this new policy. In that case, the argument would not be fallacious.

A version of the genetic fallacy is called **appeal to the person** or, its Latin name, **ad hominem**. Someone commits this fallacy when she or he argues that a claim should be rejected because of irrelevant facts about the person or people who made it. An example of this fallacy would be if someone says he doesn't believe Phil was in the library studying all weekend, because Joseph is the one who claims to have seen him, and Joseph is a jerk. As before, it is important that the fact or facts cited about the person making the claim are irrelevant. It would not be fallacious to argue that John should not be trusted about what happened at the party because he is a notorious liar about that kind of thing.

A special version of appeal to the person is **tu quoque**, Latin for "you too". This fallacy is committed when it argued that a person's claim should be rejected because the person making the claim simply fails to always live and act consistently with that claim. An example would be to argue that we should not worry about reducing carbon emissions, since Tammy says we should be worried, and her car gets terrible gas mileage. In short, you commit the tu quoque fallacy if you reject a claim merely on the grounds that the person who endorses it is a hypocrite. People are imperfect, and do not always act according to their own principles. This does not mean those principles are not good.

It is important that the claim is rejected merely because the person making it fails to live up to it. The purpose of pointing out that someone does not live and act consistently with his or her claim may be that living and acting consistently with the claim is incredibly burdensome, or even not practically possible, and that this is evidence that the claim is actually bad. This is not fallacious. For instance, it would not be fallacious to object to the mayor's proposal for a 10 PM curfew every Friday night with absolutely no exceptions, on the grounds that the mayor took his kid to urgent care after 10 last Friday. The point is not simply that the mayor is a hypocrite, but that his disobeying the curfew suggests it is a bad rule.

The fallacy of **appeal to ignorance** can happen in two ways. First, a claim is rejected on the grounds that strong evidence in favor of it has not yet been found, when we should not have expected to find such evidence even if the claim were true. Second, a claim is endorsed on the grounds that strong evidence against it has not yet been found, when we should not have expected to find such evidence even if the claim were false. An example would be to argue that Jessica must not have gone to the park because no one saw her, if there were only a very small number of people there at the time and Jessica went to a secluded area.

It is not fallacious to argue that we should reject a claim because, were it true, we would reasonably have expected to find evidence for it. Nor is it fallacious to accept a claim for that reason that, were it false, we would have reasonably have expected to find evidence against it. It would not be fallacious, for instance, to conclude that there is no goblin in your closet because you checked it thoroughly (suppose goblins can't turn invisible, or anything like that).

The fallacy of **appeal to popularity** is arguing that a claim should be accepted simply because a large number of people accept it, when property evaluating the claim requires specialized knowledge. People, even large numbers of people, are fallible. Sometimes large numbers of people will believe something false. An example would be to argue that Jones is guilty of the crime because most people think he is, though not all of the evidence has been released to the public. It is not fallacious to accept a claim because most people accept it, if it is the claim that does not require specialized knowledge to adjudicate. For instance, if you are visiting an unfamiliar town and everyone you meet tells you that Jill's Sandwich Shop is better than Joe's Sandwich Shop, it is reasonable to conclude that Jill's is better.

Similar to appeal to popularity is **appeal to tradition**, which is arguing that a claim should be accepted because it has been accepted for a long time, when viable alternatives have not been considered. For example, someone might argue against legalizing a certain substance, on the grounds that it has been illegal for decades. There are numerous factors which can lead to people accepting and holding onto a claim, and not all are good reasons. Moreover, people could have good reason to believe a claim at one point, but as new evidence emerged it might eventually become reasonable for them to reject it. Fallacious appeals to tradition function as an obstacle to reevaluating our beliefs. It is not fallacious, however, to accept a claim because it has been accepted for a long time, if the reason it has gained acceptance is that other viable alternatives have been considered and rejected for good reasons.

The **appeal to unqualified authority** is arguing that claim should be accepted because some person endorses it, even though evaluating that claim requires special skills or knowledge which the person who endorsed the claim lacks. Often when someone makes a fallacious appeal to unqualified authority, the authority cited has expertise in a different topic than the one claim is about. For instance, someone might say we should support the new tax proposal because his doctor agrees with it. A medical doctor is highly educated about the human body and medicine, but not necessarily about tax policy.

The fallacy of **appeal to force** is arguing that a claim should be accepted because there will be bad consequences if you do not, where these bad consequences are independent of the truth of a claim. For instance, if someone were to argue that you should believe he is the best candidate for the job, or else he will punch you, that person would be making a fallacious appeal to force. This threat may give you some reason to say he is the best candidate, but it does not provide any reason to think he actually is the best candidate. It is not necessarily fallacious to argue that you should accept a claim because otherwise there will be bad consequences, if those consequences are a result of the claim being true. For instance, it is not fallacious to argue that we should leave the building because there is a fire.

Work Problems 14.2

Informal Fallacies Part II

The **strawman** fallacy is arguing against a view by attacking a weakened or caricatured version of the view. For example, someone might argue against a proposal to raise the minimum wage by $5 by saying that if we are going to raise the minimum wage, why not raise it by $50? And then point out that this would likely have negative consequences on businesses. But the original proposal was only to raise the minimum wage $5, and it is much less clear that this would have a significant negative impact on businesses. A strawman argument does not succeed in revealing a problem in the criticized position, since that is not the position which is actually attacked.

The fallacy of **division** is assuming that what is true for the whole of something is also true of its parts. For instance, assuming that because something is heavy, it must be made out of heavy materials. The fallacy of division covers cases in which it is assumed what is true of a group of average is true of an individual within that group. Thus, someone might argue that Ted is a millionaire because Ted is in the room and the average income of people in the room is over $1,000,000. But it is possible for the average income in the room to be over $1,000,000 even if Ted makes nothing.

The fallacy of **composition** occurs when it is assumed that what is true of the parts of something is also true of the whole. For example, assuming that since all these chemicals by themselves are safe to consume, the result of mixing them together will also be safe to consume. But what is true about something's parts need not be true about the whole thing.

The **slippery slope** fallacy occurs when someone argues that if we allow one thing, this will lead, perhaps after a number of steps, to some unwanted further result, with the conclusion being that we should not allow the first thing. But one or more of those steps are dubious. For instance, someone might object to raising the minimum wage by $2, because pretty soon people will want to raise it by $5, $10, and then $50, and that would just be too much. It just is not clear why we should believe that those who want a $2 increase in the minimum wage will later demand a $50 increase.

Slippery slope arguments often have the structure of the argument form **hypothetical syllogism** together with **modus tollens**, both of which are valid rules of inference discussed in chapter 7. In a fallacious slippery slope argument, at least one of the premises is plausibly false, and so the argument is likely unsound. If all the premises are plausibly true, however, the slope really is slippery and there is no fallacy.

The fallacy of **false dichotomy** or **false dilemma** involves arguing in a way which presumes there are less options than there really are. For instance, someone might object to a proposal to help fix the roads in your state, on the grounds that fixing all the roads would be unaffordable. The person assumes that either all or no roads get fixed, and thereby ignores the possibility that only some of the roads are fixed.

False dichotomies have the logical form of **disjunctive syllogism**, one of the rules of inference discussed in chapter 7. Disjunctive syllogism refers to the following argument form:

Either P or Q	(Premise)
Not P	(Premise)
Q	(Conclusion)

This argument form is valid. In a false dichotomy, the premise of the form "Either P or Q" is plausibly false, because there is at least one other option. Thus, the argument is likely to be unsound.

Begging the question occurs when the conclusion of the argument is presupposed by one of the premises, and for that reason the premise is controversial. Suppose, for instance, that people start to think that John is not trustworthy, and Rachel tries to defend John by saying he told her that he doesn't lie about important things. The premise, that John said he doesn't lie about important things, presupposes the truth of the conclusion, that John is trustworthy, since if John is not trustworthy there is no good reason to believe it. And since some people think John is not trustworthy, the premise is controversial.

Arguments which begs the question are often valid, meaning that if the premises are true, the conclusion will be as well. Indeed, an argument which begs the question can even be sound, if it is valid and all the premises are true. The problem with an argument which begs the question is does not establish the conclusion but presupposes it. The conclusion might well be true, but it should not be assumed to be in a context in which it is controversial.

It is important that the conclusion is controversial. Consider the following argument:

There are both lions and tigers at the zoo; so, there are lions at the zoo.

The premise presupposes the truth of the conclusion, but if all people involved agree that there are both lions and tigers at the zoo, it is not fallacious.

Equivocation occurs when an argument relies upon multiple senses of a word or expression. An example would be if someone argued that Jack must own a goat, since he has a kid. The word "kid" can mean a human child, or a young goat. In an argument which utilizes equivocation, the premises can appear to support the conclusion if the equivocal term is taken to mean the same thing in each instance it appears; but since it has different meanings, the premises do not actually support the conclusion.

The fallacy of **hasty generalization** involves drawing a conclusion about a group based on a sample which is too small or not random. An example would be concluding that you do not like a certain field of study after taking a single class in the subject. A hasty generalization is a bad enumerative induction, a type of inductive argument discussed in chapter 12. Hasty generalization is often paired with the fallacy of **cherry picking**, which is drawing a conclusion while ignoring contrary evidence.

Work Problems 14.3

Informal Fallacies Part III

A **faulty analogy** is an argument from analogy in which important differences between the analogs are ignored. For instance, someone might reason that since his cat Butterscotch and his dog Cocoa both have fur and long tails, then since Cocoa likes swimming, Butterscotch will too. In a faulty analogy, whatever similarities hold between the primary and secondary analog or analogs are not enough to outweigh the differences, so that the premises do not adequately support the conclusion. We discussed arguments from analogy in chapter 12.

The **middle ground** fallacy involves assuming that the most reasonable position to take on some topic is somewhere in the middle of a spectrum of positions. Suppose that one candidate says the community needs to raise $100,000 for local schools, and another candidate says no money needs to be raised at all. To assume that we should raise around $50,000 would be to commit the middle ground fallacy. The fact that some position is in the middle assumes that a view on one end or the other is unlikely to be correct, but this is not so: perhaps due to a rise in enrollment new teachers are needed, and $50,000 would not be enough to pay their salaries.

Often, the middle ground fallacy results from confusing what solution is really best with what solution is most expedient or practical. Sometimes, when a group of people with different goals have to come to an agreement, the different parties involved must compromise. As a result, the answer people come to will in some sense be a middle ground position. But this does not show that the adopted position is actually the best one, only that it was the one that the majority could accept.

The middle ground fallacy can also be an outcome of the **both sides** (or **all sides**) fallacy, which involves pointing out a problematic or laudable feature of two (or more) positions or groups, while ignoring relevant context of details which makes these not equivalent. For instance, someone might point out that two politicians both formerly supported a certain bill which is now widely unpopular, and on this basis say it is best to vote for some third alternative; however, one candidate genuinely agreed with the policy, and the other candidate voted for it begrudgingly in order to secure the votes needed to pass some other bill. The **both sides** fallacy is a version of faulty analogy.

The **gambler's fallacy** involves concluding from the fact that some possible outcome has not occurred yet or recently, that it is more likely to occur soon, even though the probability of that outcome occurring is independent of the previous outcomes. For instance, someone on a catch-and-release fishing trip might reason that because she or he has caught nothing but bass so far, a bluegill will surely take the bait very soon. If the different events are independent, any one possible outcome has the same probability of occurring during each instance: assuming you practice catch and release, and assuming your fishing does not bother the fish, your probability of catching a bluegill is the same with every cast.

The **sunk cost fallacy** involves concluding that you ought to continue with something because you have already invested a great deal of resources into it, when either the desired outcome is highly unlikely, or any benefits from the desired outcome would fail to outweigh the costs. For instance, Tom has visited several stores looking for his favorite brand of soda to bring to the party, but none of them have had it in stock. Having looked for so long, Tom reasons he might as well keep looking. But if he does not go now he is likely to miss most of the party. The reason people commit the sunk cost fallacy is that they implicitly presume that if they case to pursue the goal, all the effort that went into it will be wasted. The mistake is failing to adequately recognize that continuing to pursue the goal may just lead to even greater loss, so that it is better to stop now.

The **fallacy of accident** occurs when a rule or principle which applies in most cases is assumed to apply in all; in other words, the fallacy of accident involves ignoring an exception to a principle which holds generally but not universally. For example, someone might conclude that a platypus is not a mammal because mammals do not lay eggs, and platypuses do. But while the vast majority of mammals give birth to live young as opposed to laying eggs, this is not a requirement for being a mammal.

The **Perfect Solution** (or **Nirvana**) **Fallacy** involves objecting to a proposal because it is imperfect, even if the proposal has significant positive features. For example, someone might object to a plan to cut the local poverty rate in half, since many people will still be in poverty afterwards. Although the plan does not eliminate poverty in the area, it is still extremely beneficial. Especially when dealing with complex problems, it can be exceedingly difficult to find a solution with no flaws at all. A good but imperfect solution may be superior to a solution which is flawless but unavailable.

Finally, the **fallacy fallacy** involves concluding that the conclusion of an argument is false, because that argument is fallacious. The fact that an argument contains a fallacy only shows that its premises do not provide good reason to believe the conclusion. The conclusion might, nevertheless, still be true. In showing that an argument for a claim is fallacious, you will have shown that the person making the argument has failed to provide you with a good reason to believe the conclusion. If you think the conclusion is false, you need to provide further reasons to support this contention.

Work Problems 14.4

Chapter 15. Rhetorical Trickeries, Fallacious Reasoners, and Tricksters

The first section of this chapter, following the introduction, will cover what we will call rhetorical trickeries. The sections to follow will cover how to deal with people who use fallacious reasoning and people who use rhetorical trickeries.

Rhetorical Trickeries

A rhetorical trickery is anything a person might do when participating in discussion which has, is likely to have, or is intended to have the effect of impeding good critical thinking, but which does not technically count as a fallacy since it does not necessarily involve making an argument, in the technical sense used in this book. As with fallacies, it is useful to be aware of different forms of rhetorical trickery, so that you can be aware of them. As with fallacies, there are numerous things someone can do which can undermine good critical thinking. We will only discuss a few here.

To start with, any fallacy can be used as a rhetorical trickery. This occurs when someone who knows an argument is fallacious nonetheless uses it with the purpose of compelling the audience to accept some conclusion. By seriously trying to convince people of a claim using reasoning she or he knows to be fallacious, such a person reveals her or himself to be more interested in getting others to believe or do something, than to discover what is actually most reasonable to believe or do.

That being said, some fallacies have special variants which are always rhetorical trickeries. **Rhetorical appeal to the person** (or **rhetorical ad hominem**) involves attempting to disparage an opponent or the opponent's position by mentioning irrelevant facts about the opponent, such as her or his looks, or odd habits or mannerisms. What distinguishes this from the version of appeal to the person discussed in the previous chapter is that the attack need not be part of an argument: the opponent is simply disparaged and the audience perhaps slightly more inclined to see them less favorably, even if that disparagement is not explicitly offered as a reason to think less of the opponent or the opponent's positions.

Similarly, **rhetorical tu quoque** involves distracting from criticisms of oneself by pointing out flaws in one's opponent, without necessarily explicitly presenting this as a reason that the criticisms are unfounded or should be waived. This is related to **whataboutism**, which involves distracting from criticisms of oneself or one's allies by turning attention to problematic or controversial things said and done by others, especially those in opposition to oneself. The fact that others are imperfect and have committed immoral, unlawful, or unreasonable acts does not mean that the criticisms leveled against you are illegitimate, and that you do not need to answer for them.

Another rhetorical trickery was briefly discussed all the way back in the first chapter: **appeal to emotion**. This occurs when a speaker or writer tries to rile up the audience emotionally in order to get them to be more inclined to accept a proposal, rather than offer good reasons for accepting it. A version which is often effective is **appeal to fear**, in which the speaker or writer works to build up the audience members' fear of some threat. Often, the speaker or writer presents him or herself, or an associate, as uniquely positioned to help deal with the threat. Another effective version is what might be called **appeal to optimism** or **appeal to hope**, in which the speaker or writer tries to build up the audience's excitement about the bright future that could be awaiting them, with the oft unspoken caveat that they need to support that person, an associate, or a certain proposal.

To be clear, producing an emotional reaction in your audience is not necessarily bad or deceptive. Emotional responses, even quite strong ones, are appropriate in certain situations. It makes sense to try to move your audience to be angered at injustice and enthused about pursuing a better world. When someone is trying to move an audience emotionally, try to identify what this person wants the audience to do or believe, and what reasons might be offered to do or believe it.

Appeal to fear can be effectively paired with **scapegoating**. Scapegoating is caricaturing a person or group as dangerous, malevolent, or otherwise a source of ill. Appeals to fear are often most effective when the audience is convinced they have an enemy. Scapegoating is made easier by **emphasizing extreme cases**, which involves placing excessive emphasis on particular instances of bad things happening, such as particularly heinous crimes, as a way of casting a whole group as threatening. The gorier the details, even if merely suggested, the better. It is important to understand that for nearly any group of large size, at least some group members will have committed crimes, even violent ones, and committed immoral acts. But these instances need not indicate a general pattern, nor reflect upon the nature of members of the group in general.

One way a writer or speaker may try to get an audience to be more inclined to accept certain claims is by **sprinkling poison**. This occurs when the writer or speaker makes lots of perfectly acceptable, even useful or insightful claims, while occasionally inserting a more controversial claim. The fact that the majority of the claims being made are socially acceptable or helpful inclines the audience to think of this person as credible, and a source of guidance and insight. This in turn can make the audience more receptive to this person's other, more controversial, claims. For example, a speaker might encourage people to work hard and strive to better themselves, which is often sensible advice, while occasionally disparaging government welfare programs for the poor. While it is often a good idea to take initiative and try to solve your own problems, some people may genuinely need assistance. Remember, a person can be right about some things and wrong about others, so make sure you use good critical thinking for every important independent claim.

The rhetorical trickeries discussed so far serve the purpose of helping to persuade the audience. Other rhetorical trickeries are primarily used in debates and directed at bewildering your opponent, with the persuasion of the audience being a frequent but indirect result. Someone may also use these trickeries in personal interactions in order to manipulate through confusion.

One example of this is **loaded** (or **complex**) **question**. A loaded question is a question which presupposes the truth of some other, typically embarrassing or ribald claim. For instance, the question, "Does your spouse know you're cheating?" presupposes that you are, in fact, cheating on your spouse. Such questions often leave the opponent confused and unable to effectively respond.

Another technique of this kind is **gaslighting**, which involves persistent misdirection, contradiction, deception, and lying to create confusion and to cause someone to question his or her perception of reality. We expect people to be generally honest, because most people typically are. So, when someone frequently says or suggests conflicting things, it is easy to become confused, and to think you are misunderstanding or misremembering.

Crazymaking is presenting someone with two alternatives and criticizing them no matter which they pick. In effect, it is sitting them up to fail. For instance, someone might say that you can decide where to eat lunch and suggest two restaurants. Whichever you pick will be the wrong choice, and you will be chastised for picking it. Often, the person engaging in crazymaking will treat your confusion or anger as a sign that you are being ridiculous and out of control. Gaslighting and crazymaking can be used together. Whatever view you accuse the gaslighter as having, that person can point to an occasion in which she or he said the opposite.

Finally, some rhetorical trickeries can function directly both to persuade the audience and to bewilder the opponent.

Overloading is presenting your opponent with a very large number of supposed facts, data points, and arguments in support of a claim, often in quick succession. It does not matter if many of these facts, data points, and arguments are of dubious merit. This can be very effective when done in a context in which the opponent has a limited time to respond. Even if there is enough time to respond, so much is presented that it is exceedingly difficult to clearly remember every point in order to respond to it. Given time, the opponent could respond to most, if not all of the claims and arguments being made. However, insufficient time is available. Furthermore, the opponent is worried about forgetting or mischaracterizing some of the points. As a result, she or he is unsure even of how to begin. For the audience, it can appear that lots and lots of reasons have been offered in favor of the claim. Even if the audience members recognize that many or most individual points made are not convincing, it may appear that the claims collectively constitute a powerful case for the point being made due to their large number.

The last rhetorical trickery we will discuss we will call the **implicature trap**. The implicature trap involves making a less controversial claim in a context which suggests a more controversial claim. For instance, if asked whether members of a minority religion make good employees, someone might say “Well, they have different beliefs that the rest of us.” The claim that members of a minority religion have different beliefs than members of other religious groups is, of course, likely to be true. In the context, though, this could be taken to suggest those differences in belief could lead to conflicts in the workplace. If the opponent attacks what the person literally says, she or he criticizes a rather uncontroversial claim, and appears ridiculous. If the opponent attacks the claim being suggested, the speaker can deny: “I didn’t say that!” Because the opponent is unable to effectively respond, members of the audience may be inclined to think that a strong point was made.

Work Problems 15.1

For each of the following passages, identify the rhetorical trickery being employed.

1. Theresa: What do you think about the proposal to cut welfare programs for impoverished families?

   George: I think it’s a sign of good character to take responsibility for your actions.
2. You all know I support a citywide curfew for people under the age of 18. Just last week, at midnight a teenager went into a gas station to rob it. He ended up shooting the cashier and a bystander. It was tragic.
3. Do your constituents know about your attempts to start a secret world government?
4. Is everyone here having a good time? You can do better than that. Is everyone having a good time? Well you should be. And I want you to know that you’re about to start having a *great* time, because there are great things coming. It’s time to get excited.
5. I just don’t understand John. A few nights after our date, he texted me saying he had a great time and can’t wait to go out with me again. But then the next day when I texted him asking if he wanted to do something, he said I wasn’t giving him enough space. It’s so confusing.

Dealing with Fallacious Reasoners

Having now discussed several fallacious forms of reasoning and several rhetorical tricks, it is worth discussing how to respond to someone who uses them. When engaging with such a person, it can be helpful to determine which of two possible scenarios holds. First, the person genuinely believes she or he is offering good reasons for some claim and does not realize this is not so. Second, the person recognizes that she or he is not offering good reasons to accept a claim and utilizes fallacies and/or trickeries intentionally. We will discuss dealing with people in the first category here, and those in the second category in the next section.

Sometimes a person might sincerely offer a fallacious argument for a claim, without realizing it is fallacious. Other times a person might utilize a rhetorical trickery without realizing that what she or he is doing can undermine good critical thinking. This could be because this person read or listened to someone else use the rhetorical trick, was persuaded, and then assumed that the person must have made a good case by confusing causes of belief with justification for belief, which we distinguished all the way back in the first chapter.

If such a situation arises, what do you do? Well, you might judge the belief to be innocuous enough that it does not matter. This is risky, though, since a person taken in by fallacious reasoning in one case may be taken in again in other cases, where the claim is less innocuous. Here we are confronted with an incredibly difficult question: Under what cases should we allow people to accept claims which we think there is good reason to conclude are false without challenging them to change their beliefs? We will not try to answer this question here. In any case, let us assume that you deem it worth trying to persuade the other person in this particular case.

It is important to keep in mind that, so far at least, it seems that your interlocutor is genuinely interested in finding what is most reasonable to believe. After all, that person believes something because she or he thinks there is good reason to believe it. Because of this, it is good to treat this person like a partner rather than as an opponent. You may disagree about some things, but you are both engaged in the project of trying to discover what we have the best reason to believe.

Keeping the above in mind, the best way to deal with someone who uses fallacious reasoning is, courteously, to explain the fallacy. This does not mean simply asserting that your interlocutor has committed such and such fallacy. Unless that person is already familiar with fallacies, she or he will likely not know what you're talking about and as a result might become confused. To someone who does not know much about fallacious reasoning, this can come off as abrasive, demeaning, or accusatory. You interlocutor might even think you are trying to trick them!

Instead, it may be better to start by saying that you aren't sure that person's reasons really support the conclusion. Then say that the argument seems a lot like another argument. Then offer a different example of the same fallacy, but an example which you think your interlocutor will recognize as being flawed. If necessary, help the other person to see that the two arguments are of the same kind, though they may have different content. After you have done this, and if you find it useful, free feel to share the name of the fallacy. Remember, though, it is more important that the other person not be taken in by fallacious reasoning than that she or he knows what various fallacies are called.

Remember, these fallacies have been given names because people are convinced by them often enough that it has been decided that they are worth understanding. That means that an interlocutor who has been taken in by fallacious reasoning is making an understandable mistake, one that lots of people have made and will make. In your discussion, you should show that person respect: it is easier to admit you were wrong if you can do so in a way that allows you to keep your dignity.

Alternatively, suppose you are conversing with someone who has become inclined towards accepting a claim, largely on the basis of some rhetorical trick, but by your best judgment this person is not intent on being deceptive or manipulative. As with the person who is genuinely taken in by fallacious reasoning, this person likely is interested in finding what we have the best reasons to believe. It is good to approach your interactions with this person with this in mind. It is also important to recognize that this person's critical thinking abilities have likely been compromised, especially if your discussion takes place shortly after your interlocutor was exposed to the rhetorical trick.

When conversing with such a person, your primary goal is to break the spell, to dispel the cloud of confusion brought about by the rhetorical trick. One thing that can be helpful is to explain what the trickster was doing. As with dealing with a fallacious reasoner, simply declaring that the person has used such and such rhetorical trick may not be very effective, and it could even be counterproductive. Instead, you should describe what the trickster was doing, and explain how it can manipulate the audience. If you simply assert that your interlocutor has been fooled, she or he might become defensive, and become even harder to reason with. Rather, you want to try to coax the other person into seeing the trick for her or himself.

Dealing with Tricksters

The topic of this section will be how to deal with people who intentionally utilize fallacious arguments or rhetorical tricks for the purposes of persuasion, confusion, or manipulation. It is important to recognize that, unlike the person who has been fooled by fallacious arguments or rhetorical tricks, the person who intentionally uses these tricks is not primarily motivated by the goal of discovering what is most reasonable to believe. This person is not primarily interested in learning about new ideas or understanding different perspectives. This person's primary goal is to win. In this context, winning is getting what you want. Winning might just mean having made your interlocutor look or feel foolish, or it may mean garnering the most applause during a debate. Winning could mean getting the most votes during an election or selling more books. Almost anything could count as winning, as long as it can be understood as benefiting that person in some way.

The trickster is not primarily concerned with the truth. At least, this person is not concerned with offering legitimate justifications for accepting her or his views. Because of this, if you engage with the trickster as a partner in the project of discovering what is most reasonable to believe, you will probably not make much progress. If you do so in the presence of an audience, you may even inadvertently aid the trickster's attempts to persuade others.

If you find yourself in a discussion with a trickster, explain what she or he is doing. Let this person know that these techniques are not going to persuade you. The trickster needs to understand that until she or he clearly expresses the intention of having an open and sincere discussion, you will not give this person any credence. If there is an audience present, explain the what this person is doing for the sake of the audience. Spoil the trick in order to coax your audience members to think carefully and critically about what your opponent says.

Be wary of attempts by your opponent to bewilder you. Especially if there is an audience, a person who relies on rhetorical tricks may try to bait you into doing or saying certain things. Do not take the bait. If the person asks a loaded question, explain what claim is being presupposed, and say you want to have an honest discussion without deception.

If your opponent engages in gaslighting, say that person needs to indicate her or his real position or there is no point in having a discussion. Then, if appropriate, focus on offering one or two of the best points you have in favor of your own position.

If the other person engages in overloading, explain the purpose of this tactic to your audience, and then rather than respond to the points offered directly, make one or two of the strongest points you can against your opponent's position. Afterwards, if time allows, offer your opponent the opportunity to present one or two of her or his best arguments. If your opponent engages in overloading again, explain that she or he is not actually interested in having a serious discussion, and for that reason this person should not be taken seriously.

If your opponent uses the implicature trap, ask that person to say clearly what she or he really means. Ask if your opponent intends the controversial claim which is suggested. If this person denies having said that, say you understand but ask if that is what was *meant*. Often, part of the reason for using the implicature trap is that the person does not want to explicitly endorse the view, because this could invite criticism and ridicule. Try to force your opponent to own the position. The explicitly stated but less controversial claim likely does not fully answer the question asked or resolve the problem at hand. If this is the case, draw attention to it. Force your opponent to either explicitly endorse the more controversial claim, or to back away from the topic.

Finally, be wary when someone who engages in rhetorical tricks demands a public forum. People who use rhetorical trickeries may seek to participate in public debates. Doing so suggests that this person is interested in an open discussion of views, which may seem like a good environment for arriving at the truth. But the trickster is not actually interested in an open discussion of views. Rather, this person is interested in finding an audience, and like a parasite, leeching from whatever authority the debate venue appears to bestow upon her or him to help spread the message.

The idea that the best way to discover the truth is through open debate is incredibly tempting, but you should be cautious of it. Disagreements in science are primarily resolved in the field, making measurements, or in the laboratory, doing experiments. In science and philosophy, debates occur primarily through books and articles published in peer reviewed journals. When face to face debates do occur, they take place before an audience of experts with specialized knowledge in training.

The average person may have little experience dealing with fallacious reasoning or rhetorical tricks. She or he may lack the specialized knowledge needed to distinguish claims which are well supported and claims which are not. If such a person witnesses a debate where one of the debaters is an effective user of rhetorical tricks, she or he may very well become more inclined to accept claims which are in fact not well supported.

None of this is to be taken to mean that controversial ideas should not be discussed. Indeed, they should be discussed, because through discussion the weaknesses of bad ideas can be exposed. However, you want to discuss these ideas in ways which avoid providing opportunities for tricksters. If you can, engage with controversial ideas directly, rather than debating the tricksters who peddle them. If there must be a public debate, try to find someone who is an expert in the topic and is familiar with the kind of rhetorical tricks the opponent is fond of using. The trickster is an enemy of good reasoning, act accordingly.

Part V. Appendixes

Appendix 1. Averages

The word “average” can mean either of four different things. First, sometimes the word “average” is used to mean the same as “typical,” as in “The average person enjoys listening to music.” Along with this informal use of “average,” there are three formal uses: the **mean**, the **median**, and the **mode**. The **mean** of a set of values is the sum of those values divided by the number of those values. For instance, the mean of 5, 7, 13, 26, 7, 9, 6, and 18 is:

$$\frac{5+7+13+26+7+9+6}{8}=\frac{91}{8}=11.375$$

The **median** of a set of values is the value which falls in the middle if the members are ordered by value. Thus, the median of these numbers is 7:

5 6 7 7 9 13 26.

If the number of values is even, the median is the mean of the two middle values. The **mode** of a set of values is the value which appears the greatest number of times in the set. Hence, the mode of these values is 7, since 7 appears twice and all other values appear just once.

In our example, the median and the mode are the same. Obviously, this need not be the case. For instance, for the numbers

1, 2, 6, 8, 8, 9, 12 15, 23, 23, 23, 42, 75,

the median is 12 and the mode is 23. Just as well, neither the median nor the mode need be the same as the mean, which in this case is

$$\frac{1+2+6+8+8+9+12+15+23+23+23+42+75}{13}=\frac{247}{13}=19$$

Also clear is that neither the mean, the median, nor the mode must be typical, and so none of these formal senses of “average” is the same, or even reliably approximates, the informal sense.

Appendix 2. Variance and Standard Deviation

The **dispersion** of a set of values refers to how far the values in that set are from the mean for that set. While there are many ways of measuring dispersion, two of the most common are **variance** and **standard deviation**. Here we will explain how these are calculated.

Let the mean of a set of values be represented by the lower-case Greek letter mu: μ. Let the number of values in the set be represented by the lower-case letter n. Let the value of each member of the set be represented by the x_i. Here, "i" is a variable for a number: x_1 represents the value of the first member, x_2 of the second, x_3 of the third, and so on. For each value x_i, we find difference of x_i and μ, and then square it. Thus:

$(x_1 - \mu)^2$

$(x_2 - \mu)^2$

$(x_3 - \mu)^2$

and so on.

Next, we find the sum of each of these differences. That is

$(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2 + \ldots$

This can be symbolized more simply like so:

$$\sum_{i=1}^{n}(x_i - \mu)^2$$

This method of symbolization may seem intimidating if you are not used to it. It is useful as a very compact way of writing what could otherwise be a very long formula. The symbol "Σ" is just the upper-case Greek letter sigma. In mathematics, Σ is used to indicate that you are to find a sum. What values are you supposed to sum? Well, the answer to each formula

$(x_i - \mu)^2$,

for each value of i. What are the values for i? Well, “i = 1” below Σ tells us we take the value from every member i of a certain set. Which set? The set represented by “n” above Σ. In this case, n = the number of values in the set. Thus, for each member of the set, we find the difference between that value and μ, square the result, and then add all those together.

Finally, we divide the result by the number of items in the set. Thus, the formula for variance looks like this:

$$Variance = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

The standard deviation, which is standardly symbolized by the lower-case Greek letter sigma, σ, is just the square root of the variance. So,

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

Now that we know how to find the variance and the standard deviation, let’s discuss what they mean. We will just focus on variance, as the relation between variance and standard deviation is obvious.

Obviously, for any given value we plug into i, the formula

$(x_i – \mu)$

tells us how far that value is form the average. Since the product of any number and itself is positive, by squaring this formula,

$(x_i – \mu)^{2}$,,

we ensure that the distance between the values for i and μ is measured positively. Thus,

$$\sum_{i=1}^{n}(x_i - \mu)^2$$

measures the total (squared) distance of all values in the set from the mean.

Since n is the number of values in the set,

$$\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

measures the mean of the (squared) distances of all values in the set from the mean. Finally,

$$\sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

compacts this value.

Recall that μ stands for a mean, and σ for a standard deviation. The range of values which fall between

$\mu - \sigma, \mu + \sigma$

are within *one standard deviation* from the mean. The range of values which fall between

$\mu - 2\sigma, \mu + 2\sigma$

are within *two standard deviations* of the mean. The range of values which fall between

$\mu - 3\sigma, \mu - 3\sigma$

are within *three standard deviations* of the mean. And so on.

It turns out that approximately 68% of members of the population will fall within 1 standard deviation, approximately 95% fall within 2 standard deviations, and approximately 99% fall within 3 standard deviations. Where X is a member of the population, we have:

$\Pr(\mu - \sigma \leq X \leq \mu + \sigma) \approx .683$

$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx .955$

$\Pr(\mu - 3\sigma \leq X \leq \mu - 3\sigma) \approx .997$