THINKING WELL - A LOGIC AND CRITICAL THINKING TEXTBOOK

Andrew Lavin
Butte College

LibreTexts™

Thinking Well - A Logic And Critical Thinking Textbook (Lavin)

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 07/29/2025

TABLE OF CONTENTS

Detailed Licensing

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

About the Author and Book

About teh book

Authored and Edited by Andrew Lavin with pages from:

- Matthew J. Van Cleave's Introduction to Logic and Critical Thinking
- Matthew Knachel's Fundamental Methods of Logic
- Jason Southworth and Chris Swoyer's Critical Reasoning: A User's Manual
- Inserts on Truth and Intellectual Vices by Michael Fitzpatrick

This text made possible by an Open Educational Resources Grant through Butte College and the Textbook Affordability Project at CSU, Chico.

About the authors

This textbook's first edition was funded by grants from Butte College and CSU, Chico. Since then, I have continued to put unfunded work into improving it. This text is free and
always will be, but if you are using it and enjoying it, a donation would be much appreciated and would help fund future improvements to this free resource. Please consider a donation of somewhere around $5 from each reader to support the project. To reiterate, though, this is a free textbook and no donation is necessary.

My Teaching Resources and Portfolio can be view here: https://alav.in/

Places to send donations:

- Paypal: andrewlavin@live.com
- Venmo: @Lavin-Andrew

CHAPTER OVERVIEW

1: Basic Concepts

The most important thing we do as human beings is learn how to think. This is *important* in two senses of the word: it's important to human beings because it is the most distinctively unique fact about our species—we think rationally and abstractly—but it's also important because it the most wide reaching capacity we have—it touches virtually all aspects of our lives. Having a heart that pumps blood or a body capable of certain physical activities might be more *fundamental* meaning more crucial to simply surviving, but *thinking* underlies a broad range of activities without which we would be living less than full human lives.

The common title of this course is "Logic and Critical Thinking." So, we can think about the course as having two main components: the study of formal logic and the study of the tools and strategies of critical thinking. This text is structured in a bit of a "sandwich". Units on critical thinking and then formal logic, and then units on more critical thinking topics.

First, Logic. We'll define logic more fully later, but for now: logic is a sort of reasoning that is mathematical in its precision and proofs. It's like math with words and concepts, in a sense.

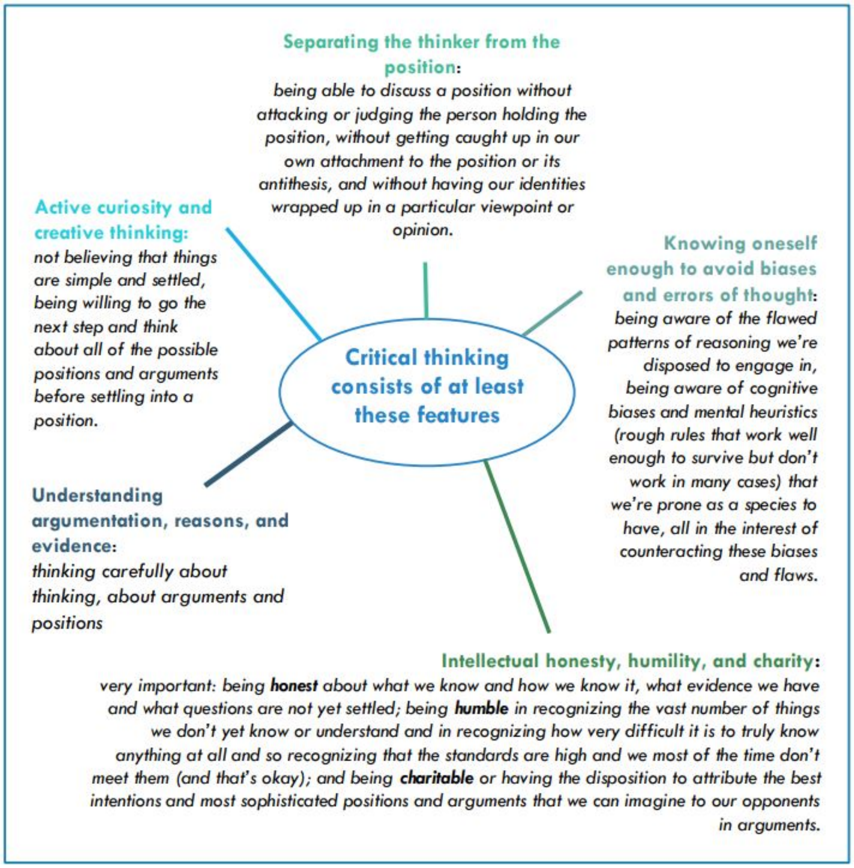
Oh no! Not math! I'm no good at math.

Don't worry, dear student. Logic is more straightforward than a lot of the complex concepts that get discussed in math classes. Even better, all of logic can be broken down into simple, step-by-step processes that a computer can do. You just need to follow the steps carefully and you'll be guaranteed the right answer every time. There's no magic to it, no special skills or abilities needed. You just need to follow directions carefully and put a bit of work into it.

Next, let's get a bit of a definition of critical thinking going. Critical thinking is primarily the ability to think carefully about thinking and reasoning—to have the ability to *criticize* your own reasoning. 'Criticize' here isn't meant in the sense of being mean or talking down or making fun of. Instead, I mean the word in the sense of, for example, how a coach might take a critical stance toward her players' skills—he throws high every time, she doesn't lead with her foot, they ride too forward in the saddle, etc. 'Critical' here means something more like 'reflective' or 'careful' or 'attention to potential errors'.

So to engage in critical thinking is to engage in self-critical, self-reflective, self-aware thinking and reasoning—thinking and reasoning aimed at self-improvement, at truth, and at careful, deliberate, proper patterns of reasoning.

There are many definitions of what critical thinking is, but here're my thoughts:



As you can see, being a critical thinker involves training yourself to have a lot of good habits and dispositions. It involves developing rational virtues so that when the time comes to think about something complex, you are naturally disposed to think well. It doesn't happen overnight and it certainly doesn't come for free—no one is born with it. We all need to train ourselves and educate ourselves to stay guarded against errors in reasoning.

1.1: Vital Course Concepts

Have you ever been in a conversation with someone and found that they were mischaracterizing what you were saying, bombarding you with seemingly unrelated information, giving up on reasoning at all and instead simply blustering? It feels like you want to figure out what is true and really try to understand the issue you're discussing, but the person you're talking to just wants to make some talking points or get a few shots in before plugging their ears and singing to themselves. Like having the discussion, to them, is just a game that they want to win at any cost. I've been there, loads of times. Not fun.

Rhetoric vs. Reasoning

In ancient Greece one of the most famous Western philosophers of all time—Socrates—noticed that there were different ways that people came to believe what they believe.

- Some people listened to the poets, the oracles, the playwrights, etc. for truth. They looked to ***Tradition*** to find out what they should believe, how they should act, what reality was like, etc. Think Homer: they read *the Iliad* and *the Odyssey* and took life lessons from its pages.
- Others found that by manipulating those around them they could come to win arguments reliably. They found that discussions were sort of like jousting matches in that the person who won the argument was usually the person who was the quickest wit and had the best tools at their disposal. They didn't care about what was true, or good, or right. They only cared about winning and about ***Rhetoric***. They were the lawyers and politicians of their day (not to say that there aren't good, honest, principled lawyers and politicians). They were called *Sophists*.
- Socrates himself preferred a different route he called ***Dialectic***, where a person made a claim and then others asked critical questions about their claim or position until they found themselves confused and unsure of what to believe. This process of breaking down our presuppositions helped us move past *false confidence*. The closer we move toward what was called *aporia*—a state of "impasse" where we're not sure which direction to take—the more we're able to consider what the truth is about from all of the baggage we carry from our childhood, from previous conversations, from popular media, etc.

Sophists are interested in good *rhetoric*. Good rhetoric is important. In fact, we have whole courses and sometimes whole departments devoted to good rhetoric. We need to find ways of expressing our ideas in a way that gets the right reaction from our intended audiences.

Philosophers, like Socrates, are interested in good *reasoning* or *argumentation*. We philosophers (and I count myself among them) care about rhetoric only to the extent that we're not misleading our audiences or turning them off of our position without good reason. We want to explore what the best reasons are for accepting or rejecting different positions. We don't want to win an argument if we're wrong. We're ready and willing to revise our beliefs if we find out that we don't have good reason for accepting them.

With this in mind, we can posit a distinction between rhetoric and reasoning. Two things you'll be worried about to different degrees in different situations. Sometimes you'll care a lot about how your audience will receive your arguments and other times you'll care more about simply getting it right.

> **Rhetoric**: Is primarily concerned with the impact of an argument or piece of writing or speech or the like. How effectively is it producing the effect I want in my audience?
>
> **Reasoning**: is concerned with insight, discovery, truth, and understanding. The goal isn't to produce a certain impact in the audience, but instead to collectively discover what the best position on a given question is or what the objective merits and demerits of an argument or position are.

This is a course about argumentation and reasoning, we'll be interested in Rhetoric only insofar as it gets in the way of good, honest, clear argumentation. (That isn't to say that rhetoric *always* gets in the way. Sometimes, in fact, good rhetoric can amplify good reasoning. Good reasoning is dry and inaccessible without good rhetoric!)

Propositions

So we're interested in how *arguments* work. What makes them tick? What makes the good ones good and the bad ones bad? How can I make a series of statements and then think that I've "proven" or "demonstrated" a further statement? In order to understand arguments, we'll have to start with the fundamental building blocks: propositions or statements:

Definition: Propositions

Propositions are statements that can be true or false. *This is **the** fundamental concept of the course. Take the time to understand it clearly.*

If a statement can be true or false, then it's a proposition. Note that a sentence and a proposition aren't the same thing. Not all sentences are propositions.

When we reason, we make statements or consider statements and then we back those statements up with reasons and evidence, draw out the implications and consequences of those statements and so on. There's a technical distinction between a statement and a proposition, but we will use them interchangeably here. For our purposes, a statement and a proposition are the same thing.

Some sentences don't express propositions at all. This means that they can't be true or false. You can't disagree with them, you can't argue about whether they're right or wrong, you can't question them. Not because they're indubitable (un-doubt-able), but simply because it wouldn't make any sense to disagree with them!

If I said, "Can we please go out to dinner tonight?" you can't respond with disagreement, saying "I don't know about that claim, it doesn't sound right to me." I haven't made a statement, so you can't say I've stated something false. Similarly, if you say "wash your hands before dinner" I can't respond with "that's false." It wouldn't make any sense. These types of sentences don't express propositions. They're non-propositions.

Definition: Non-Propositions

Sentences that aren't statements about matters of fact (or fiction). They don't make a claim that can be true or false. They:

- **Exhort**: Let's go get drinks! Let us go hiking on Tuesday!
- **Command**: Go to the store later to buy me some cheese. Don't do that.
- **Plead/Request**: Would you please stop that? Please read me a bedtime story!
- **Question**: What is the capital of the UAE? How much do the pineapples cost?
- **Perform**: I hereby adjourn this meeting! I pronounce you husband and wife!

Complex Propositions

Okay, back to propositions. Sometimes propositions are simple, and sometimes they are complex. Meaning sometimes they can be broken down into simpler propositions and sometimes they're already as simple as they could be.

Definition: Simple Propositions

Simple Propositions have *no internal logical structure*, meaning whether they are true or false doesn't depend on whether a part of them is true or false. They are simply true or false on their own.

- The GDP of the United States is $5.
- The Sky is Blue.
- Freedom should be the highest value of the state for its citizens.
- Harry Potter wears glasses.

Definition: Complex Propositions

Complex Propositions have internal logical structure, meaning they are composed of simple propositions. Whether they are true or false depends on whether their *parts* are true or false.

- The GDP of the United States is **either** $5 **or** it is $12.
- *True if **the GDP is $5** or if **the GDP is $12***
- The Sky is Blue, **but** it doesn't look blue to me right now.
- *True if **the Sky is blue** and if **it doesn't look blue to me right now**.*
- **If** freedom should be the highest value of the state for its citizens**, then** we should promote it in our laws and policies.

- *True if it can't be that "**freedom should be the highest value of the state for its citizens**" is true while "**we should promote freedom in our laws and policies**" is false.*

In short, each proposition is either a simple proposition or a combination of simple propositions. Simple propositions are true or false just based on how the world is, whereas complex propositions are true or false just based on whether or not the simple propositions that make them up are true or false.

I am an elephant

...is a false proposition if I say or think it. It's false because of the way the world is—I am not in fact an elephant.

I am an elephant or I am a human

...is a true complex proposition if I say or think it. The way "or" propositions work is that only one of the simple propositions needs to be true. The left proposition "I am an elephant" is false, but the right one "I am a human" is true. So, since complex propositions depend on their parts for their truth values, the complex proposition as a whole is true!

We can learn to break propositions down into parts. This is an important skill to grasp so that you can understand all of the separate claims someone is making in a single sentence. People often make a host of claims in a single sentence, and you'll want to be able to separate them.

Breaking down complex propositions usually involves identifying the little sentences that make up a complex sentence. So instead of saying "Bobby doesn't want to play basketball, but he does want to play videogames." I notice that the "but" connects two independent thought: Bobby doesn't want to play basketball. Bobby wants to play videogames.

"Either you know everything there is to know, or I'm a monkey's uncle and you're not as smart as I thought you were." Breaks down into three separate propositions since there's an "either...or..." and also an "and".

Breaking down Propositions: separate out the statements that can be independently true or false. It's a bit tricky and interpretive, but we're just trying to grasp the basic concept here.

We'll get into this more later in the course, but for now it's good to have some facility with the basic idea: some propositions don't have parts that can be true or false independently, while others do. We use words like 'and', 'or', 'Either...or...', 'but', and 'if...then...' to identify multiple independent propositions.

Example 1.1.1

- Marcos is taking four courses this semester and working in his parents' store 20 hours a week.
  - Marcos is taking four courses this semester.
  - Marcos is working in his parents' store 20 hour a week this semester.
- Frankie, Johnny, and Luigi went to dinner
  - Frankie went to dinner.
  - Johnny went to dinner.
  - Luigi went to dinner.
- Karen is smart but not very motivated to do well in school or to try to find a job that uses her talents.
  - Karen is smart
  - Karen is not very motivated to do well in school.
  - Karen is not very motivated to try to find a job that uses her talents.

Okay, so now we know that there's an important difference between sentences which express propositions and those that do not. We also know that some sentences express multiple simple propositions and some express only one simple proposition. We've hopefully got a good grasp on what a proposition is at this point.

Inferences or Arguments

The topic of the course is the *argument* or *inference*. What's that?

Definition: Inference or Argument

An **Inference or Argument** is any purportedly rational movement from evidence or premises to a conclusion.

Any time you're being asked to accept one claim on the basis of or because of any number of other claims, you've got an inference/argument. "I believe x, because of y, z, and w" or "Because a, b and c, we have to believe that d."

The following is from Knachel, Fundamental Methods of Logic, CC-BY 4.0 Int'l

If we're reasoning by making claims and backing them up with reasons, then the claim that's being backed up is the conclusion of an argument; the reasons given to support it are the argument's premises. If we're reasoning by drawing an inference from a set of statements, then the inference we draw is the conclusion of an argument, and the statements from which its drawn are the premises.

We include the parenthetical hedge—"supposed to be"—in the definition to make room for bad arguments. Remember, in Logic, we're evaluating reasoning. Arguments can be good or bad, logically correct or incorrect. A bad argument, very roughly speaking, is one where the premises fail to support the conclusion; a good argument's premises actually do support the conclusion.

To support the conclusion means, again very roughly, to give one good reasons for believing it. This highlights the *rhetorical purpose* of arguments: we use arguments when we're disputing controversial issues; they aim to persuade people, to convince them to believe their conclusion.[3] As we said, in logic, we don't judge arguments based on whether or not they succeed in this goal—there are logically bad arguments that are nevertheless quite persuasive. Rather, the logical enterprise is to identify the kinds of reasons that *ought* to be persuasive (even if they sometimes aren't).

So you've got some support for a conclusion and then that conclusion. The relationship between that support and that conclusion is supposed to be *rational*—we're supposed to believe that the support we're given *proves* or *demonstrates* or *gives us reason to believe* the conclusion.

Less abstractly, here's an example:

Example 1.1.2

Bob Marley wrote "One Love"

Bob Marley sang the best rendition of "Don't Worry, Be Happy"

Bob Marley wrote "Three Little Birds"

Bob Marley wrote "No Woman No Cry"

Bob Marley wrote "Buffalo Soldier"

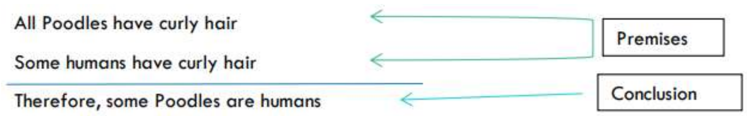
So Bob Marley is the greatest musician of all time

We're being asked to believe a number of things here. First, we're supposed to believe that Marley wrote each of these songs. We're also being asked to believe that his version of "Don't Worry, Be Happy" is the best version ever. Finally, and most importantly, we're being asked to believe that *because* all of these things are true, it follows that Bob Marley is the greatest musician of all time.

This isn’t a very good argument. It’s not being very good has *nothing to do with the conclusion*. I love Bob Marley and I do think he has written some of the best songs of all time, but I don’t think that these premises entail this conclusion. That is, even if we accept all of these premise, we need not accept the conclusion. “Oh yeah?” someone can reasonably reply, “those are all amazing songs, yes. And I don’t dispute that Marley wrote them, but none of them is as good as Bohemian Rhapsody. Queen is therefore better and Bob Marley cannot be the greatest musician of all time.” There’s no inconsistency with believing that this argument uses good premises to support a possibly-true conclusion, but doesn’t really demonstrate that conclusion using those premise. We can believe the conclusion is true without accepting that the argument supports the conclusion.

Premises and Conclusions

All arguments have a common anatomy:



There may be many premises, but they’re all supposed to be statements (propositions) which support or demonstrate the conclusion—whether directly or indirectly. A premise is a proposition which lends credence to the conclusion. It’s supposed to be a group of statements that, if you accept that they’re true, make it the case that you rationally *must* (or, weaker, *should)* accept the conclusion. That’s not the case above in our Bob Marley argument. Here’s an argument where it is true:

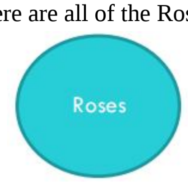
All Roses are Red

All Red Things are Ugly

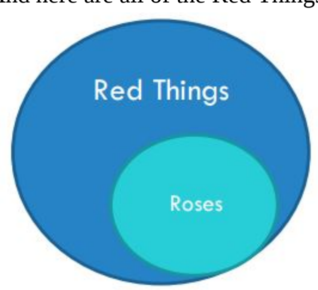
All Roses are Ugly

If it is in fact true that all roses are red and if it is in fact true that all red things are ugly, then it follows *with absolute certainty* that all roses are ugly. Try to accept the premises and reject the conclusion. Can’t do it. It’s impossible.

Here are all of the Roses:

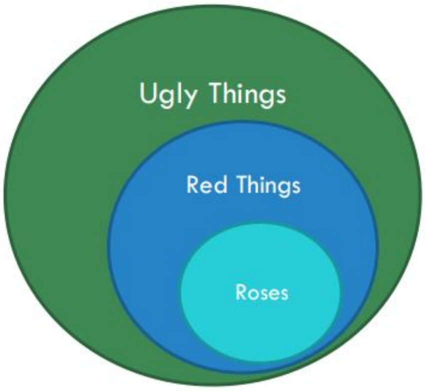


And here are all of the Red Things:



According to the first premise, all of the roses are red things. Notice how all of the roses fit inside all of the red things? That’s a graphical way of representing the claim that there are no roses outside of the category of red things. If you’re looking for a rose, you’re looking for a red thing. No non-red roses. All roses are red.

Here are all of the Ugly Things:



According to the second premise, all of the red things are ugly things. See how they all fit inside the circle? Well, if the roses are in the red things circle and the red things circle is in the ugly things circle, it follows that the roses circle is in the ugly things circle. That make sense?

Notice how we probably don't want to accept the conclusion here. Most of us are on board with roses. We think that roses are beautiful or at least that they're not ugly. So we don't like this argument. **But we can't, when rejecting an argument, reject the conclusion directly.** Why? Well, because the premises are supposed to be *proving* the conclusion. If the premises are true, then the conclusion must be true. We have to, therefore, reject one of the premises here. The second premise is clearly false. There are lots of pretty and red things. Red roses for instance. Furthermore, not all roses are red! The two premises were false in this case. If, however, you thought both premises were true, you'd have to accept the conclusion as well.

Imagine your two friends are dating. If you want to invite one of them to hang out in a group setting, both of them will generally want to come. All of a sudden, they're a package deal, right?

They have a relationship with one another such that if you take one, you have to take the other as well. That's sort of how premises and conclusions work. They have a logical relationship with one another such that if you think the premises are true, you *must* also think that the conclusion is true. They're a package deal.

The previous argument illustrates the point that arguments sometimes are *bad*. There are sometimes reasons to reject an argument (mind you, not to decide that its conclusion is false, but instead to decide that it didn't demonstrate its conclusion).

Arguments can go wrong in only two ways:

1. **Bad Inferential Structure**: every argument with the same structure as this argument is bad (invalid or weak). The premises don't in fact demonstrate or maybe even support the conclusion. In other words: we can accept the premises as true without being compelled to accept the conclusion. There's something wrong with this argument's general structure.

2. **False Premise(s)**: this particular argument has a premise/assumption that is false. There's something wrong with this argument's particular content.

The following is from Matthew J. Van Cleave's *Introduction to Logic and Critical Thinking*, version 1.4, pp. 2-3. Creative Commons Attribution 4.0 International License.

So, to reiterate: all arguments are composed of premises and conclusions, which are both types of statements. The premises of the argument provide a reason for thinking that the conclusion is true. And arguments typically involve more than one premise. A standard way of capturing the structure of an argument is by numbering the premises and conclusion. For example, recall Sally's argument against abortion:

> ✔ Example 1.1.3
>
> Abortion is morally wrong because it is wrong to take the life of an innocent human being, and a fetus is an innocent human being.

We could capture the structure of that argument like this:

1. It is morally wrong to take the life of an innocent human being

2. A fetus is an innocent human being

3. Therefore, abortion is morally wrong

By convention, the last numbered statement (also denoted by the "therefore") is the conclusion and the earlier numbered statements are the premises. This is what we call putting an argument into standard argument form. We can now give a more precise definition of an argument. An argument is a set of statements, some of which (the premises) attempt to provide a reason for thinking that some other statement (the conclusion) is true. Although arguments are typically given in order to convince or persuade someone of the conclusion, the argument itself is independent of one's attempt to use it to convince or persuade. For example, I have just given you this argument not in an attempt to convince you that abortion is morally wrong, but as an illustration of what an argument is. Later on in this chapter and in this book we will learn some techniques of *evaluating* arguments, but for now the goal is to learn to *identify* an argument, including its premises and conclusion(s). It is important to be able to identify arguments and understand their structure, whether or not you agree with conclusion of the argument.

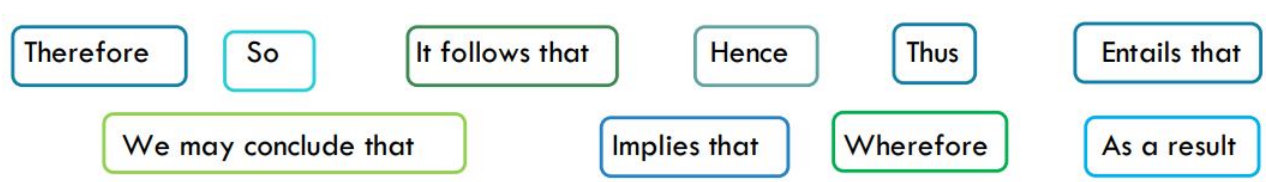
How do we identify Premises and Conclusions? Good question! First, we can sometimes identify premises and conclusions simply by recognizing the role they play in an argument. Here's an argument, for example:

> ✔ Example 1.1.4
>
> *Migratory butterflies are facing strain or possible extinction due to the overdevelopment of lands along their migration routes. In developing the routes, humans have tended to remove milkweed, which is a central food source for migratory butterflies.*

There are two claims here. One seems to support the other one and not the other way around. Which one seems like the claim being supported here? Which one seems like it's doing the supporting? Good! Right-o, chum. See how the first sentence raises a question and the second sentence answers it? "What's the relationship between developing land and migratory butterflies?" or simply "Why should we believe that?" The second question answers these questions: we should believe the first sentence because developing means removing milkweed, which is a food source for butterflies.

Second, we can recognize conclusions and premises by identifying certain words being used: these are called **conclusion indicators** and **premise indicators**. If one of these words is used, typically that means that you've spotted a conclusion or a premise (depending on the indicator).

Conclusion Indicators all have the general sense of "I've told you some things or I'm about to tell you some things, now here's what I want you to believe." They have a conclusive feel to them. Here are some especially common ones:



Premise indicators, on the other hand, have the general sense of "from this fact I'm going to infer something else". Here are some common premise indicators:

Here's an example argument that I've packed with Indicators:

Example 1.1.5

__In that__ the legislature has not approved it, and __given that__ it is unconstitutional for me to do it on my own, __I must conclude that__ there is no legal way for me to complete the project using only executive orders and the budgetary authority given to the executive branch. Furthermore, __as indicated by__ the general lack of public support for the plan, __it follows that__ I will be acting in line with the popular will on this issue. __Therefore,__ I must not allocate money to make Fridays "free pizza days" __since__ to do so would be a great abuse of executive power.

Factual Claims and Inferential Claims

Each argument makes two different sorts of claims, as we saw with the Bob Marley example above. There are a number of independent factual claims: claims about what is in fact the case or about what the world is like. There is also a somewhat hidden claim that the premises presented compel us to accept the conclusion presented—that there's a good inference from these premises to that conclusion.

Facts and Inferences

argument makes two sets of claims:

- The **Factual claims** are in the premises: the arguer is claiming that all of the premises are true.
- The **Inferential claim** is often implied. The arguer is also claiming that the premises give conclusive support to the conclusion.

Remember that these are the **two ways an argument can break down**: it can make a false factual claim (a premise can be false) or its premises can fail to support its conclusion (the implicit inferential claim can be false).

So since we couldn't possibly find an argument with the same general structure as the previous argument that has true premises, but a false conclusion, we conclude that the structure of that argument was deductively valid. We went through the steps to show that the argument is valid. We demonstrated it with the circles above (these are called "Euler[1] Diagrams"). Here's the structure without the particular content:

All Ωs are Δs

All Δs are Ψs

All Ωs are Ψs

So now we have the same *types* of propositions in the same order, but we're no longer just talking about roses and red things and ugly things. Any argument with this form will be valid: you won't be able to reject the conclusion if you accept the premises. This is what it means to have good inferential structure if you're a deductive argument. We'll get into the difference between deductive and inductive arguments later.

Here's an *invalid* inferential structure:

All Ωs are Δs

All Ψs are Δs

All Ωs are Ψs

How do we know it's invalid? Well, easy. We find an example of an argument with the same argument form such that the premises are true and the conclusion is false. How about this example?

Example 1.1.6

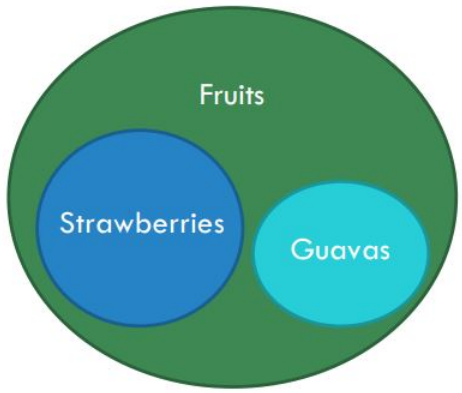
All Guavas are Fruits

All Strawberries are Fruits

All Guavas are Strawberries

Guavas and Strawberries are two different sorts of fruit, but that doesn't mean that they're the same thing!

Let's check out what the Euler Diagram looks like:



See how strawberries and guavas don't need to overlap *at all* to make the premises true? But the conclusion says that they *totally overlap*. So the premises do not entail the conclusion. We can accept the premises without accepting the conclusion. We only need this one counterexample to show that this argument structure is invalid.

What does it mean to have a false premise? Pretty simple. Each premise makes a statement about how the world is. The world either is or isn't that way, so each premise either is or isn't true. Any argument with a false premise isn't a good argument. All of the premises must be true for an argument to have successfully demonstrated its conclusion.

Definition: Parts of any argument

The **Conclusion** is the claim that the whole argument is intended to support or demonstrate or prove. It's the reason we make an argument: to support or demonstrate the conclusion.

The **Premises** are the claims, evidence, ideas, etc. that are intended to support the conclusion

- They are the assumptions we are asked to take on board. If they are true, then the conclusion either must be or is likely true as well.

Here's a heuristic or rough rule that can sometimes help you identify premises and conclusions. If you are told a premise, you'll likely not understand *why* you were told it until you see that it fits into an argument. If you are told a conclusion, you'll likely wonder what reason the person has for thinking it's true. The conclusion is the "point" of bringing up the premises: to demonstrate that the conclusion is true. The premises are reasons we have for believing the conclusion.

Premises make you ask: Okay, but what does that mean? What's the point? Why tell me that?	**Conclusions make you ask:** Okay, but why would anyone believe that? Give me reasons or evidence for accepting that claim.
• **Hey, did you know that the US spends more on researching a new fighter jet than it would on making college tuition free for everyone?** • *Okay... where are you going with this?*	• **The Portland Trail Blazers are by far the more cohesive basketball team in the NBA this season.** • *Wait, why do you say that? What's your evidence?*

On Truth

> *The following section was written by Michael Fitzpatrick*

A guiding assumption of this textbook is that truth is our aim when evaluating particular arguments and their conclusions. But what is truth, and why is it so important? After all, many people today seem to think that truth is just whatever a person believes, or that truth doesn't matter as much as other values such as economic success, political pragmatism, or self-fulfillment. Contrary to these trends, this textbook affirms that thinking well flows out of the nature and value of truth, and that the value of truth should guide how we think.

The name 'truth' gathers together our human concerns about accuracy and sincerity, to use the two underlying notions proposed by Bernard Williams in his book *Truth and Truthfulness*. When asserting propositions, adopting beliefs, creating blueprints, drawing maps, giving directions, following cooking recipes, or even just trying to see how much taller a teenager on a growth spurt is since last year, what we're aiming at is an accurate understanding of our world and our projects in the world. Accuracy is a *more or less* concept, one we apply for instance in the game of darts. When throwing at a dart board, we aim for the bull's-eye, and we are more or less accurate depending on how close we get to the target. But truth is not simply about accuracy; it also concerns sincerity, that dimension of representing ourselves and others and our world in ways that are genuine, faithful, trustworthy, and can be taken as presented. When we ask a friend for directions to the gas station, we're not just depending on the accuracy of their instructions, but also their sincerity in not wanting to deceive us and in wanting to be someone who is trustworthy.

Truth understood as the interplay of accuracy and sincerity describes a fundamental way human beings exist. To interact truthfully in the world is to experience ourselves, our neighbors, and our environment as it really is, not as we wish it to be. The mode of truth highlights our capacity to receive new information from sources other than ourselves.

Truth is, as Martin Heidegger suggests, "uncovering," the unconcealing of what is with us in the world, seen for its own sake. In his book *Kant and the Platypus*, Umberto Eco suggests that truth involves our encounter with "lines of resistance" in the world, ways in which reality pushes back against our concepts and perceptions. In a particularly pregnant passage, Eco describes reality as a continuum with definite shape,

> *As if to say that in the magma of the continuum there are lines of resistance and possibilities of flow, as in the grain of wood or marble, which make it easier to cut in one direction than in another. It is like beef or veal: in different cultures the cuts vary, and so the names of certain dishes are not always easy to translate from one language to another. And yet it would be very difficult to conceive of a cut that offered at the same moment the tip of the nose and the tail.*
>
> *If the continuum has a grain, unexpected and mysterious as it may be, then we cannot say all we want to say. Being may not be comparable to a one-way street but to a network of multilane freeways along which one can travel in more than one direction; but despite this some roads will nevertheless remain dead ends. There are things that cannot be done (or said). (53)*

Truth names the lines of resistance contouring what we can say, believe, do, experience, and imagine. Resistance does not entail impossibility; we are capable of astonishing creativity that can be used to distort reality, whether by lying to others, creating elaborate conspiracy theories, or engaging in our own self-deception. The point is not that we cannot do these things, but that to do them we have to overcome the resistance of reality. When we are sincere people seeking to aim accurately, we are letting ourselves be shaped by the "frictions" we experience as beings in the world.

The foregoing portrait of truth helps to uncover the value truth has and its indispensable role in human life. Truth is both intrinsically and instrumentally valuable: something is intrinsically valuable when it is a end in itself, and instrumentally valuable when it is a means to a end. Truth, like justice and goodness, is both. Truth is also essential to other values humans hold, as well as to a healthy psychological life. We'll cover each of these in turn.

1. Truth has intrinsic value because it satisfies our curiosity and wonder at existence. To be human is to explore and discover. We try to figure out how electrons can be both a wave and a particle; what creatures live in the darkest regions of the oceans; why whales are mammals that evolved back into the sea; and what the heck is a platypus anyway?! We even investigate the mystery of our selves—from our ingrained irrationality to our capacity for language to our religious tendencies to the enigma of consciousness, we humans want to understand who we are. As we explore our own identities and values, we discover what is true about us, which means we transcend our own fantasies, illusions, and folk stories about who we are.

2. Truth also has instrumental value for protecting ourselves against manipulation by other people. Consider this: you are not free to believe whatever you want. What I mean is, you can't just will yourself to believe something. Try it – will yourself right now to believe that there is a pink elephant in the room with you. No matter how hard you try, you can't just decide to believe it (at least not *sincerely*; you could say you believe it and not mean it). But you can decide to get others to believe something that is false or dangerous. You can trick or manipulate them into believing falsehoods, as a practical joke or to take advantage of them. But if you can do this to others ... they can do it to you too. If our beliefs are not guided by what is true, then we are vulnerable to other people influencing our beliefs according to what they *want* us to believe. Politicians do it all the time.

3. Truth is essential to being a responsible, ethical human being. We are not just physical beings governed by the laws of physics; we have responsibilities to ourselves and others. We need to care for our younger siblings, pay our credit card bills on time, not cheat on our sweethearts, and stand up for those less fortunate than ourselves. But meeting these responsibilities requires a reasonably accurate portrait of the world. If we don't know the truth about other people—their needs and their hopes and their fears—or the truth about the material situation we all find ourselves in, then we can't meet our responsibilities. Knowing what my sister needs from me requires knowing my sister, knowing the truth not only about her needs but how best to meet them given the resources I have. Action presupposes truth.

4. Finally, truth is essential to a healthy human psychological life. Bernard Williams, in his essay "Deciding to believe," describes a man who knows his son is dead but does not want to believe it. Suppose the man decides to undertake some act of self-deception so that he no longer believes his son is dead. The problem is that there are other true beliefs the man probably has about the world that imply his son is dead (for instance, that his son never sends letters or comes for a visit). Williams writes,

> *The man gets rid of this belief about his son, and then there is some belief which strongly implies that his son is dead, and that has to be got rid of. Then there is sanother belief which could lead his thoughts in the undesired direction, and that has to be got rid of. It might be that a project of this kind tended in the end to involve total destruction of the world of reality, to lead to paranoia. Perhaps this is one reason why we have a strongly internalised objection to it. If we are not going to destroy all the evidence—all consciousness of the evidence—we have to have a project for steering ourselves through the world so as to avoid the embarrassing evidence. That sort of project is the project of the man who is deceiving himself, and he must really know what is true; for if he did not really know what was true, he would not be able to steer around the contrary and conflicting evidence. (151, from Problems of the Self)*

The attempt to deceive ourselves into believing one false belief seems to lead to a life of real paranoia and psychological breakdown, as well as the incoherence of needing to know what is true in order to avoid what is true. Of course, believing true beliefs can also lead to psychological upheaval. The realization that Nazi Germany had engaged in extensive crimes against humanity and undertook a holocaust against Jewish and disabled peoples was enormously difficult for many German citizens to accept, and required tremendous revision in their beliefs about themselves, their communities, and their cultural identity. Yet upheaval by true beliefs is externally motivated, coming not from an internal paranoia or avoidance of reality, but from the "lines of resistance" and "uncovering" of reality we described above on the nature of truth. However difficult those truths may be to accept, they are more likely to lead to psychological stability in the long-term than, say, a holocaust denier who has to spend the rest of their lives refuting evidence and testimony about what really happened.

There are more reasons beyond these to value the pursuit of truth, but hopefully this provides a sense of why this class is focused on how to reason as well as humanly possible so as to live along the grain of the universe.

The Principle of Charity

Okay, Andrew's back. Now let's talk about a really important habit to get into. A sort of norm for reasoning well:

> *Always interpret your opponent/interlocuter's position or argument so as to make it as strong or defensible as possible.*

There are three reasons for this: one having to do with our goals in having reasoned discussions; another having to do with simple strategy if you are indeed interested in winning a debate; and finally one moral reason for following the principle of charity.

If you're interested not in *winning*, but in *understanding*, then of course it doesn't help to argue against the weakest version of someone's position or the weakest justification available for someone's position. For instance, if you want to *understand* the moral issue of abortion, then arguing with someone who makes a super weak version of an argument for or against abortion rights won't really help you understand the issues at play in the moral debate. You might win the debate on that day, but you won't have understood the issue with any more clarity.

When you disarm your conversant by letting them know that you understand their position and why someone might believe it, you open the door to more honest and open dialogue that allows for more understanding of each other's viewpoints.

Even if you are interested in winning and you just want the most effective strategy for winning a debate, the principle of charity is still your best bet. Here's an example of what not to do:

✓ Example 1.1.7

My opponent has argued against the idea that immigration is a fundamental human right. She must mean that even amnesty-seekers don't have the moral right to immigrate away from immediate threats to life and limb. That position is totally ridiculous.

This isn't very interesting. When you're arguing, you want your opponent to be the hardest version of themselves to critique so that when you do critique them, your critique is the most interesting critique available. Think about how much more interesting it is if someone actually bolsters their opponent's position by providing justifications for their position and then showing that their position is *still* wrong. That's the kind of debate take-down I want to see.

✓ Example 1.1.8

My opponent has argued against immigration as a fundamental human right by appeal to simple scarcity: there's not enough to go around. This, I'm afraid is simply false. There is more than enough to go around if we're willing to redistribute resources effectively. Nevermind that, though, since there's a stronger justification for my opponent's position: that states have the fundamental right of sovereignty, which includes controlling traffic across their borders. This, we might think, is essential to what it means to be a state. This is a very interesting argument, but it still fails to convince me. Even if states have a right to border regulation, it doesn't follow that individual human beings don't still have a right to immigrate to where the greatest promise of prosperity is.

Isn't this interesting? Wouldn't you rather be in dialogue with this person or listen to a debate they're in than someone who only attacks the weakest interpretation of their opponent's position?

You don't want your critique to be against a straw figure version of their argument (the easiest-to-refute version) because all they have to do is revise their position slightly and they can side-step your critique. You've set them up to make your critique null and void by simply clarifying their position as the stronger version.

This is also a good principle for living your life. You want to always attribute the most virtuous intentions to others and to the actions of others. This makes us easier to be around, more fun to converse with, and more empathetic and understanding people. It's therefore a moral imperative that we treat each other with charity. After all, don't you want to be given the benefit of the doubt?

We want to interpret each other's actions and arguments as being as rational as possible so that we are in the best standing rationally speaking. We want to attribute the best, most rational intentions allowable by our evidence to those around us so that if we have a problem with what they're doing, we at least have given them the benefit of the doubt and are more likely to correctly characterize what they're up to. We want to ascribe the most defensible and reasonable arguments and claims to people with whom we disagree because we certainly don't want to spend our time critiquing an argument or position *that isn't theirs*! We want to focus our energy on the best position or argument they have available to them *because we are interested in finding out what the best thing to believe is*—what's true. We aren't interested in winning for the sake of winning.

---

[1] Pronounced "Oiler" because it's German.

1.2: Kinds of Inferences

Let's look at two different inferences:

Inference i:

The sign says only 3 more miles to the coast, I suppose we're getting close!

***Inference d*:**

The definition of a scab is a union member who works during a strike, Manny is a union member who is working during a strike, so Manny is a scab.

Notice how even if we accept the premise of inference i, we need not accept the conclusion. Ten gazillion different things could make the sign inaccurate. Maybe the coast moved due to erosion or seismic activity, maybe the sign was stolen from its intended location and moved 10 miles inland, maybe the sign was a practical joke in the first place. Who knows?

For inference d, though, it's not so open-ended. We have a definition and the claim that an individual meets that definition. If the definition of *x* is *d* and *a* is *d*, then *a* is an *x*. If we disagree with the conclusion, we either have to reject the definition, or the description of Manny. We can't add new information to change the conclusion. Even if Manny is a good guy, or an alien in disguise, or a really pro-union guy, or supporting three kids and a wife with cancer, he's still a scab if that is in fact the definition of a scab and if that is in fact a true description of Manny. Harsh, but *certainly* true if the premises are true.

The point of comparing inferences I and d is to see that there are two fundamentally different sorts of inference. We call an inference **inductive** if the support the premises provide for the conclusion is less than certain—if the premises don't *guarantee* the conclusion. We call an inference **deductive** if the premises provide conclusive support for the conclusion—if they guarantee the conclusion or make the conclusion *certain*.

Deductive arguments are mathematical arguments like proofs and the like, logical arguments, arguments from definition, etc. If the premises are true and the argumentative structure is good, then the conclusion *must* be true.

Inductive arguments are arguments from analogy, arguments from qualified authority, causal inferences, scientific hypothetical reasoning, extrapolations from samples, and so on. Even if the argumentative structure is great, the truth of the premises only even makes the conclusion *probably true* at best.

There's a third kind of argument where we select the best explanation from all of the available plausible explanations. We won't spend time on it, but it's worth noting its existence. It's sometimes called "abduction."

Definition: Kinds of Inference

Deduction: arguments where the premises *guarantee* or *necessitate* the conclusion

- Mathematical Arguments, Logical Arguments, Arguments from Definition

Induction: arguments where the premises make the conclusion *probable*.

- Analogies, Authority, Causal Inferences, Scientific Reasoning, Extrapolations, etc.

Inference to the Best Explanation or Abduction: arguments where the best available explanation is chosen as the correct explanation.

Validity

Remember that *truth* is a property of propositions. That is, only propositions can be true or false. **Arguments can never be true or false**. It simply doesn't make any sense to claim that an argument is true or false.

Okay, let's talk about deductive arguments for a hot minute. Deductive argumentative structures are either valid or invalid. An invalid argument structure is one where the premises don't guarantee the truth of the conclusion, but they *should*, given the type of argument involved. For instance, if it's a mathematical argument, then it's premises *should* guarantee its conclusion so it's deductive. But if its premises *don't in fact* guarantee its conclusion, then its an invalid deductive argument.

A valid argument structure is an argument structure where the premises guarantee the conclusion. That is, if the premises are true, the conclusion follows *necessarily*. It's impossible for the premises to be true and the conclusion false. If 2+2=3, and 6-3=3, then *necessarily*, *beyond any doubt*, 2+2=6-3. No ifs, ands, or buts about it. It's *impossible* for that conclusion to be false without at least one of those premises being false as well. All true premises and a valid argument means the conclusion *must* be true.

Keep in mind that validity is about *structures*. So the previous paragraph's arithmetical argument has the structure a+b=c, d+e=c, therefore a+b=d+e. Anything we sub in for the letters, if we create two true premises, will necessitate a true conclusion.

What's a valid argument that has true premises? That's called a sound argument. Soundness is about both structure and truth: you have to have a good structure and true premises to be a sound argument. An unsound argument, conversely, is an argument that either is invalid or has at least one false premise.

Truth, Validity, Soundness

What's **Truth**? A proposition makes a statement about the world and the world either is or isn't the way the proposition describes it to be. One proposition claims that the Gross Domestic Product of the United States of America is approximately $14 Trillion. To find out whether this is true or false, go figure out what the GDP of the US is. Is it approximately $14 Trillion? Another proposition claims that there's a brown cat on the front porch of your house. Is this true? To find out, just go look at the world: is there in fact a cat on your front porch? Is it a brown cat?

The propositions that make up an argument (the premises and the conclusion) are all either true or false. As with all things in philosophy, there is a lot more to say about the complexities here. Some of the earliest philosophy in the Western philosophical tradition is philosophy of logic or language. Aristotle, for instance, asked whether it's true or false that there will be a sea battle tomorrow. Isn't it contingent? Aren't there lots of indeterminant factors involved in determining whether or not a sea battle will in fact take place? If so, it seems like that proposition is neither true nor false yet. So not every proposition is either true or false. We need not, though, deal with such issues. We can proceed as if every proposition is determinately either true or false.

Remember: *Validity* is a property of argument structure: it means "this structure is such that if the premises of any argument with this structure are true, then the conclusion of that argument *must* be true."

It means: arguments of this structure will never have all true premises and a false conclusion. The structure *guarantees* the truth of the conclusion given the truth of the premises. Almost like the structure carries the truth of the premises directly to the conclusion without fail. A reliable one-way transporter of truth.

A sound argument is an argument that has a valid structure but then also has true premises. If an argument is sound, and if validity means the conclusion *must* be true if the premises are true, then the conclusion must be true, then what do we know about the truth of the conclusion of any sound argument? Yes! You're so smart: the conclusion of any sound argument is guaranteed to be true.

Definition: Truth, Validity, Soundness

Truth: *propositions* are either true or false

Validity: good deductive *argument structure*: True premises make the conclusion necessarily true. (if not, it's an Invalid structure)

Soundness: Valid deductive argument, all True premises. (If not, it's an Unsound argument)

$$\boxed{\text{All True Premises}} + \boxed{\text{Valid Structure}} = \boxed{\text{Sound Argument}}$$

Truth, Strength, Cogency

Switching over to **inductive** arguments, we find an analogous set of properties. Again, inductive arguments are made up of propositions, which can be true or false.

The biggest difference is that even good inductive arguments only offer probabilistic support for their conclusions. Meaning accepting all of the premises doesn't *necessitate* that one accept the conclusion, it merely gives one more or less strong reason for accepting the conclusion. So the argumentative structure of an inductive argument isn't either good or bad, it's a matter of degree (and often a matter of what the actual content is). An inductive argument can therefore offer *stronger* or *weaker* inductive support for its conclusion.

"Cogent" and "uncogent" are the words we use in place of "sound" and "unsound" for inductive arguments since inductive arguments cannot be sound or unsound. Cogent, therefore, means all true premises and the premises give strong inductive support for the conclusion.

Consider these two arguments:

I saw a black cat *Therefore all cats are black*	*I saw the Sun rise in the East every day of my life and everyone I know reports the same and history books and ancient astronomers report the same, so the Sun will rise in the East tomorrow.*

Notice how the argument on the left provides pretty weak support for the conclusion. I can believe that the speaker in fact saw a black cat and still think that's a bad reason for concluding that all cats are black. The right argument, though, is much stronger. There's much more evidence and the nature of the evidence makes the conclusion much more probable given the truth of all of the premises.

Definition: Truth, Strength, Cogency

Truth: *propositions* are either true or false

Strength: Inductive *argument*: true premises make conclusion probably true.

Cogency: Strong inductive argument, all True premises.

$$\boxed{\text{All True Premises}} + \boxed{\text{Strong Inductive Support}} = \boxed{\text{Cogent Argument}}$$

Again, inductive arguments are collections of propositions (the premises and the conclusion(s) are all propositions. And each of these propositions might be either true or false depending on whether it accurately describes reality.

Note

An inductive argument cannot be valid. Why? Because a valid argument **guarantees** the truth of the conclusion. But an inductive argument only justifies its conclusion **to some level of probability**.

1.3: Chapter 1 - Key Terms

- **Rhetoric**
- **Reasoning**
- **Propositions**
- **Non-Propositions**
- **Inferences/Arguments**
- **Non-Inferences**
- **Premises**
- **Conclusions**
- **Premise/Conclusion Indicators**
- **Deduction**
- **Induction**
- **Truth**
- **Soundness**
- **Validity**
- **Strength**
- **Cogency**

1.E: Chapter One (Exercises)

Exercise $1.E.1$: Rhetoric vs. Reasoning

Is each primarily engaged in rhetoric or primarily engaged in reasoning?

A. We're going to invade that country, and we're going to win. The enemies of democracy will find no quarter on the face of the Earth. The mighty arm of justice will smite those who fight in the name of evil and tyranny.

B. We're going to invade that country, and here's why: the threat the current government poses to the American people is clear and present and furthermore the massacres taking place through the military are just short of genocide. We have, I think, a moral duty to intervene to prevent the further loss of life—American or Foreign.

C. My client is innocent of these trumped up charges and furthermore is the victim of a relentless campaign of defamation motivated by the most cynical and dare I say racist motives this country has seen in a prosecution. Any decent person can see that though my client is no angel, she is nevertheless no murderer.

D. My client is innocent of the charges brought before you in this case. I'll lay out why in what follows. She has an air tight alibi: she was at her brother's funeral at the time of the murder. She has no motive: she was taking piano lessons from the victim at the time, but the victim's logs show that they had a good and healthy rapport. They were even friends on Facebook and those records indicate a friendly relationship.

Exercise $1.E.2$: Propositions vs. Non-Propositions

Is each of the following a proposition or a non-proposition? (remember that propositions are sometimes false)

A. Don't eat my food when I've clearly labeled it.

B. Please try to be more considerate.

C. The moon is made of cheese.

D. The moon is the most awesome thing in the universe.

E. The Betelgeuse Nebula is approximately 2 light years from Earth.

F. How can you act that way when you know it hurts her feelings?

G. Nobody knows what lives on the dark side of the moon.

H. Stop talking about the moon.

I. The dark side of the moon isn't really dark.

J. Pink Floyd is one of the greatest musical groups of all time.

K. Yaaaaah! Pink Floyd!!!!!

L. For real, though, please stop talking about the moon.

Exercise $1.E.3$: Breaking down complex propositions

Break down each complex proposition into its component simple propositions. The words 'and', 'but', 'or', 'if...then...', and 'if and only if' tell us (at least at this introductory stage) we're dealing with a complex proposition and also tell us what the simple propositions are.

A. Y'all don't know nothing, but at least I know something that'll help us get out of this situation.

B. I've always been a self-reliant person, and I don't generally like to travel together, but if you can get me to Chicago, then I'd be mighty grateful.

C. If I had a cow, then I would have all the milk I needed.

D. Either Valeria knew what she was doing or she was completely ignorant of the effects that her actions were likely to have.

E. Robots are either sentient or they are just machines.

F. An ear of corn is either used for animal feed or for applications in human cuisine.

G. Samir, Raj, and Asia were late to school today.

Exercise 1.E.4: Argument or Non-Argument?

Identify whether each is an argument or inference, or whether it is instead a set of disconnected statements, an explanation of *why something happened the way it did*, an account or story, or something else that doesn't have an inferential connection in it.

A. She only cheated on the exam because her financial aid depended on it.

B. We know she cheated on her exam because the instructor caught her stealing glances at her neighbor's test and found notes hidden in the sleeve of her jacket.

C. He only cheated on his boyfriend because his boyfriend was neglecting him.

D. Rabies is an often-deadly viral infection that causes hydrophobia or the inability to drink water. There are treatments available, but they must be administered quickly.

E. You shouldn't go around spreading stories about people since you wouldn't want people doing the same to you.

F. No one has ever been to the moon, so we don't know that it *isn't* made of cheese.

G. "Boy when you walk by every night talking sweet and looking fine, I get kinda hectic inside." – Mariah Carey

H. "I never sleep, 'cause sleep is the cousin of death" – Nas

I. No one should ever sleep, because sleep is the cousin of death

J. Every time I walk by here, you're sitting around doing nothing. I guess you don't have anything to do with your time.

K. John Lennon was shot and killed. So was Bobby Kennedy and his brother John. Dr. King and Malcolm were killed around the same time.

L. "Hey Jude. Don't make it bad. Take a sad song and make it better." – Paul McCartney

M. "(you shouldn't) carry the weight of the world on your shoulders, for you know that it's a fool who plays it cool." (Hey Jude again)

N. We need to reverse our course when it comes to greenhouse gas emissions. If we don't, the Earth will soon become inhospitable to human life. We must act now.

O. As Joni Mitchell said, we shouldn't pave paradise and put up a parking lot. To do so would be to make the all-too-common mistake of failing to recognize what we've got until it is gone.

Exercise 1.E.5: Premise and Conclusion Indicators

For each proposition you find, identify the premises and the conclusions based on the indicators. If there are no indicators for a proposition, try to figure out if it's a conclusion or a premise.

A. Since we can't go out until the restaurant opens, and given that we were trying to save money, we should just hang out here until closer to dinner.

B. Why should we save the wetlands, you ask? Simple: because wetlands protect us during storms as they slow water flows and surges.

C. As there has never been a storm of such strength in the Florida Keys, one must conclude that there is not likely to be a storm of such strength in the Florida Keys in the years to come.

D. Nearly every student in the class scored less than 70% on the exam. This entails that the exam was too difficult to accurately assess student learning and therefore, since exams are supposed to accurately assess student learning, the exam must be revised and readministered.

E. Insofar as the losing candidate received a majority of the popular votes, and given that we live in a democracy, we may conclude that the election went against the basic principles of our country.

Exercise 1.E.6: Factual Claims and Inferential Claims

For each, determine which factual claims are being made and which inferential claims are being made. There will likely be a list of factual claims and then one (often implied) inferential claim.

A. The Affordable Care Act is in a death spiral. Premiums are getting higher, and as premiums get higher, the people will stop purchasing policies, and if the people stop purchasing policies, then the insurance companies will pull out of the exchanges, and if that happens, then the whole system collapses.

B. College isn't designed around the goal of producing good plumbers and electricians and welders. Furthermore, college is expensive and college is time-consuming. So, we shouldn't expect everyone to go to college.

C. We should pass the Affordable Care Act. There is an epidemic of chronically-ill citizens without health insurance due to their pre-existing conditions and many citizens simply can't afford health insurance.

D. No one wants to be put in the position where they are faced with a deadly intruder without the proper means to protect theirself and their family. Gun laws make it probable that someone will end up in that situation. Therefore, we can't enact gun control legislation.

Exercise 1.E.7: Charitable Interpretations

For each, supply a charitable interpretation of the argument or position. Give a sort of argument for why someone might believe this or accept this argument. Even if you don't believe the conclusion, we still need to be able to give the strongest version of the argument or position we're engaging with.

A. I think the Earth is only about 5,600 years old. The fossil records are inconclusive because carbon dating makes a lot of questionable assumptions.

B. We should keep spending more on the military every year. China is only becoming more of a threat.

C. Borders are not necessary. The concept of a closed border came about fairly recently in human history and having hard borders only creates violence and a coercive market for human trafficking.

D. Policing is inherently unjust. Communities in our society should be able to regulate themselves.

E. We need to build a colony on the Moon!

F. Aliens have certainly visited us in secret.

G. Astrology is real, my personality matches my horoscope more often than not.

CHAPTER OVERVIEW

2: Language, Meaning, and Definition

Critical thinking, on one understanding of the idea, is the ability to ask the right questions. Some of the right questions are questions about the words used in an argument or used to express a position. What do they mean? No, what, *specifically*, do they mean? When someone says something like "immigration is a problem." What do they mean by "immigration"? Are they referring to illegal immigration? Legal Immigration? All immigration? A specific nationality? A specific subset of illegal immigrants? What do they mean by "a problem"? Do they mean "we need to find out how to support these people as they struggle to survive?" or do they mean "we need to protect ourselves from these people"? We don't really know exactly what they mean until we've clarified it with them (or sometimes looked at the other things they're saying and inferred what they mean).

2.1: Breakdown of meaning

What's a concept? A concept is something we use in thought, like an idea. We combine concepts to form new concepts, we think about the world *through* concepts, and we form new concepts. For example, when we see a bunch of striped horse-like animals running around we can use the concept *zebra* to group them together and then learn more things about them. When we see a bunch of different people betting on the price of commodities, we use the concept *futures trading* to understand these disparate activities under one idea. One way of putting the point is that concepts are the things expressed by words or phrases. The phrase "futures trading" expresses a concept (an abstract meaning, or an idea in our mind). The word "zebra" expresses a different concept. I can form whole thoughts in my mind by putting concepts together to form *propositions.*

Zebras are striped

Futures trading is risky and morally suspect

Here we have the concepts ***Zebra***, ***Striped***, ***Futures Trading***, ***Risky***, *and* ***Morally Suspect****. We can recombine these to make new thoughts:*

Zebras are morally suspect

If a thing is striped, then it's risky

Futures trading is striped

These aren't very good thoughts or propositions, but they're thoughts or propositions nonetheless.

The communication of concepts (expressed using words or phrases) breaks down in many ways. That is, when we try to communicate our ideas to other people, we often mess up. We leave ambiguity or speak vaguely and therefore don't effectively communicate what we intend to communicate. Two such breakdowns are especially important to be able to distinguish: Vagueness and Ambiguity.

Vagueness

Vagueness is **problematic *unspecificity***. Meaning the concept or word or phrase admits of borderline cases (has fuzzy boundaries), and/or doesn't tell us enough to mean much of anything.

If you can't easily pinpoint where a concept start and stops—if you can't draw a clean line around all of the things the concept applies to, excluding all of the things the concept doesn't apply to—then you have a vague concept. It's not necessarily a *problem* to have a vague concept, though, unless you *need* a more sharp or specific concept.

- That man is tall!
  - 1 inch taller than average? 2 inches taller than average? 7?

- Water is “hot” just in cases where it feels much warmer than room temperature.
  - “much warmer” isn’t any more specific than “hot”
- You’re no friend of mine unless you’ve been there for me.
  - Does that mean being available when you call me? Does that mean going out of my way to comfort you? What *degree* of support are you expecting? The boundaries aren’t clear.
- We’ll do whatever is prudent to help the victims of this natural disaster.
  - Does that mean donate some money to mercy corps? Does that mean spare no expense? What *degree* of support are you offering? There are many borderline cases for which this statement won’t help us predict what support will be offered. I like to call these “weasel statements” because they allow one to weasel out of obligations later by saying “I said we’d do whatever is prudent and it turned out not to be prudent to do much of anything, so I never lied.”
- The winner is whoever dances the best.
  - Wait, what does that mean? How do we measure “goodness” of dancing? It might be better to say that the winner is whoever the judges score the highest. That’s specific and quantitative—it’s easy to measure.

- I have lots of stuff to do, so I won’t be able to make it.
  - Look, we all have lots of stuff to do. If you truly have so much stuff to do that it justifies cancelling our plans, then I’d want to know *how much* you have to do. Do you have to take the dogs for a walk? Or do you have to rewrite an entire term paper in an evening because your hard drive crashed the day before the paper was due? These are very different scenarios.

Ambiguity

Ambiguity is problematic in that the ambiguous word or phrase ***admits of multiple distinct interpretation*****s**. A sentence could be read in different ways based on how we interpret a word or phrase. There is **grammatical ambiguity**, and also **semantic ambiguity**.

Definition: Grammatical Ambiguity

Grammatical ambiguity is where the structure of a sentence (like a dangling modifier or a poorly-placed pronoun) make the sentence compatible with more than one reading.

Definition: Semantic or Lexical Ambiguity

Semantic or Lexical Ambiguity is where a word or phrase could mean multiple different things, each of which makes the sentence as a whole have very different interpretations.

- John went to the park to meet Bob, but he never arrived.
  - Pronoun ambiguity: to whom does “he” refer? Grammatical Ambiguity
- It’s difficult to walk a dog wearing a dress because of all the funny looks.
  - Who is wearing the dress? The dog? The Person? This is a dangling modifier. So it’s Grammatical ambiguity.
- I believe in freedom, so I believe the government should do nothing but protect me from threats foreign and domestic.
  - “Freedom” is ambiguous between positive freedom and negative freedom (freedom from constraint vs. freedom to accomplish one’s goals). Semantic ambiguity.
- Headline: Prostitutes appeal to Pope
  - What does “appeal” mean here? Appeal as in “attract” or appeal as in “plead with”? Semantic ambiguity.

- She critiqued them for playing soccer poorly.
  - This is a little tricky to identify, but it is a dangling modifier because it's ambiguous between "she critiqued them poorly" and "playing soccer poorly." The modifier "poorly" doesn't clearly attach to the critique or the soccer playing because it could attach to either.
- He fed her cat food
  - Did he feed a woman cat food? Did he feed a female cat cat food? Did he feed a woman's cat with food? Grammatical ambiguity.
- We all saw her duck.
  - We saw a duck that belongs to her? Or we saw her perform the action of ducking? The word "duck" is ambiguous here.

We need to be able to distinguish between these two breakdowns in meaning to fully understand each of them. Sometimes meaning breaks down because someone is trying to express a concept but uses ambiguous wording which leaves us uncertain *which* concept they're trying to communicate. Other times, someone is trying to use a concept to make a point, but the concept they choose is vague, which leaves us uncertain how to apply the concept to the concrete cases because the concept is indeterminate—we're not sure what to do with certain borderline cases.

2.2: How does meaning work? Definition and Concepts

Let's think a little bit about how meaning actually functions. How does a word like "pig" refer to actual objects in the world? What does it mean to get more or less specific? How do we determine which actions are "lying" or "cheating" and which actions are just somewhat similar but are ultimately innocent? What does it mean to put things into categories with words?

Understanding all of this a bit better can help us to understand the meaning behind the words we use a bit more clearly.

Intension and Extension

Intension

Intension is essentially the specific information that a word or phrase conveys. The more specific a word or phrase or concept, the more intension it has (the more information it gives you).

The intension of a word is basically the set of instructions it gives you for identifying what thing or things the word is referring to. So the more intension a concept or definition has, the more specific it is in that we have more information we can use to identify things that fall under the concept.

"Human Female" refers to all of the women by having meaning that applies to all and only women.

"Zebra" applies to all and only the zebras in the world, past, present, future (and maybe some zebras that are only possible, but this is complicated) by having all of the instructions contained in it necessary for us to identify which things are and are not zebras.

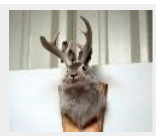
"Potato" has less intension than "fingerling potato" in that the group of potatoes *contains* the group of fingerling potatoes. There's more information in the second one than in the first, so it's *more specific*, and there's *more intension*.

Extension

Extension is the set of things referred to by a word, phrase, or concept.

The extension of a word is an actual set of things in the world, past, present, and future.

"John Claude Van Damme" refers to one specific person, whereas "Mammal" refers to a great deal of things: all of the mammals. The more extension a term has, the bigger the set of things it refers to.

Zebra	*has the extension:*	(All of the actual zebras, past, present, and future) (and maybe all the merely possible zebras too)
Jackalope	*has the extension:*	 (Only imaginary animals, but all of the jackalopes. Fred the imaginary Jackalope, Marisol the imaginary Jackalope, ...)

Increasing and *Decreasing* Intension

- Animal, mammal, feline, tiger
- Wolf, canine, mammal, animal

Mammal has more intension than animal, since you not only have to check if it's an animal, but also have to check if it's a mammal (if it has mammary glands and gives birth to live young, etc.) in order to see if it's the animal you're talking about. Still more feline is more specific and gives you more information about what specific individuals fall within the extension. Tiger is more specific still.

Wolf is very specific, canine less specific, mammal and animal are less specific still.

Decreasing and *Increasing* Extension

- Animal, mammal, canine, wolf
- Tiger, feline, mammal, animal

The group of things called ‘animal’ is pretty big. It’s a very small percentage of animals that are mammals, so the extension has gotten much smaller. Again a very small group of mammals are canines and even fewer are specifically wolves. So the extension is decreasing.

Tigers are a relatively small group compared with felines (house cats, leopards, panthers, etc). Mammals are a bigger group than felines and animals are a bigger group than mammals. So the extension is increasing here.

> Intension and Extension are typically **inversely related**: the more specific you get, the fewer things you’ll be talking about. If you want to talk about more things, you’ll have to get less specific.

This makes sense, though, right?

Things you can do with Concepts

Synthesis and Analysis are words we hear all of the time, but very seldom do we get any sort of idea what they actually mean. With that in mind, during this discussion about meaning and concepts, it seems like a good idea to think about what these words mean. They are, in short, the two different things one can do with concepts. One can build up from simpler concepts to a complex concept —this is called Synthesis. One can alternatively break down a complex concept into relatively simple concepts—this is called analysis. Synthesis is bringing different elements together to form a complex. Analysis is breaking down a complex into its elements.

- **Synthesis**: building up to more complex concepts from simpler ones
  - If you have a **state** in which *citizens elect officials*, you have a **Republic**.
  - A **saw** that **cuts** *using a chain blade* is a **chainsaw**.
  - A **dog** that **performs a service** *for someone with a disability* is a **service dog**.
- **Analysis**: breaking down to simple concepts from complex ones
  - A **Cup** is: a **thing**, that **holds liquid**, *for drinking*.
  - A **teacup** is a **cup** *used for tea*.
  - A **hammer** is a **tool** *used for pounding*.

Who Cares?

What’s the upshot of all of this talk of extension, intension, synthesis, analysis?

This is a really basic and philosophically naïve picture of how meaning works: We have concepts that can be synthesized into more specific concepts or analyzed into less specific concepts. Concepts have two aspects to them: they refer to some set of things in the world (their extension), and they do so by having a particular *meaning* or *sense* which describes all of those things it picks out. The more specific they are, the fewer things they can refer to in the world. The less specific, the more things they can refer to. The ideal definition is one that is neither too broad nor too narrow and so has the right level of specificity.

Each problem with meaning (like vagueness and ambiguity) is a problem with the meaning of a word or phrase and so is most likely a problem with either extension or intension. Most of the time with ambiguity, a single word or phrase is linked to different intensions and so by virtue of that is linked to different extensions. With vagueness, the problem is *by definition* too little specificity and so a problematic lack of intensional information. It’s helpful to “look under the hood” of what is going on with vagueness and ambiguity so we can approach concepts carefully and with a bit of understanding of how they work.

Types of Definitions

Not all definitions are created equal. Some are better than others. These are just three of the questions you might ask yourself when encountering a new definition. Understanding what sort of definition it is you are dealing with can go a long way towards being able to evaluate that definition as a basically good or basically problematic definition.

Other textbooks offer a wide variety of distinctions between definitions: enumerative vs. genusspecies vs. subclass vs synonymous, and then again theoretical vs. precising vs. ... It goes on and on. It doesn’t seem strictly necessary, though, to understand each of these different kinds of definition in order to make basically good judgments about the quality of a definition. I’ve instead boiled the lot down to three basic distinctions that seem truly useful in evaluating whether you’re dealing with a useful definition or an inaccurate or useless definition.

Stipulative or Descriptive?

First, is this definition attempting to *stipulate* new meaning? That is, is it trying to invent a new word or use an old word in a very precise or perhaps artificial way? Or maybe it's doing the same thing as a dictionary definition: just trying to describe the way the word is in fact used by speakers of the language.

Definition: Stipulative definitions

Stipulative definitions either define a new word or define a familiar word in an unnatural way for the purpose of an argument or theory.

- A "Hill House" is a house in the Hollywood hills.
- For the purposes of this study, when I say "justice" I'll mean "equal distribution."
- A "rave review" for the sake of this argument, is a review of 4 or more stars out of 5.

Definition: Descriptive definitions

Descriptive definitions are standard dictionary definitions, they attempt to describe the way a word is actually used.

- A brick is a solid rectangular object made by drying clay.
- A Bitcoin is a unit of cryptocurrency run by the Bitcoin blockchain network.
- A mother is a female primary caregiver or a bearer of children.

It would be pretty odd to open up the dictionary and find a definition like the following:

Tall: a measurement of height applying to objects the tops of which extend 6'4" or more higher than their bases.

Who put Merriam and Webster in charge of what makes a thing tall? It seems oddly specific to choose 6'4" as the objective meaning of the word "tall." So this seems like a bad descriptive definition. What about, though, if someone said:

Okay, I want all the tall people in the back row. If you're 6'4" or more, then you're tall, please go to the back row so we can take our picture.

Doesn't sound quite as unnatural to me. What do you think? In this case, it seems okay to *stipulate* that tall means 6'4" or more *in a particular context*, but it doesn't seem okay to act as if *that's just the definition of 'tall.'* So it seems okay when it's stipulative, but wouldn't be good if it was descriptive.

Too Broad or Too specific? Or Apt?

Some definitions are simply too broad. They cover too many things. For instance, consider the following definition:

Currency is anything of value.

Well, teapots have value, but I doubt very much that you would be comfortable calling a tea pot "currency" (think about calling it "money", sounds weird, doesn't it?). Horses have value, but they don't count as currency, right? So the group of things with value is much larger than the group of currencies (Yen, Mark, Rupee, Dollar, etc.), and so the definition doesn't work. And then there's the other issue of trying to understand what "value" means here. Monetary value? Exchange value? Sentimental value? This definition is too broad.

Other definitions are simply too narrow. They don't seem to cover the whole group of things they're meant to. Here's one such example:

Example *[Math Processing Error]*

A (American) liberal is anyone who spent the 60's burning bras and draft cards.

That seems to maybe cover a subset of American liberals: hippies. Even then, this might still be too narrow of a definition for hippy. Certainly, there are liberals who weren't even born yet in the 60's. There will be some who aren't even born *right now*, let alone in the 1960's. There were also liberals *before* the 60's. This definition is far too restrictive: many people in fact count as liberals that would, by this definition, be included amongst the non-liberals.

*Or maybe a definition is **apt**, meaning it captures the right group of things (the right extension) by being not too narrow (not too much intension) and not too vague or broad (not too little intension).*

Argumentative vs. Neutral

Some definitions, like those in the dictionary, are at least trying to define the term or concept in a way that doesn't slant in one way or another. Others, though, are biased towards a particular attitude or judgment or perhaps towards a particular conclusion. This is why we call these *argumentative definitions*, since they often implicitly contain an *argument*.

Definition: Argumentative definitions

Argumentative definitions have a clear ideological bias behind them. They're trying to get you to feel a certain way or make a certain moral judgment about the thing being defined. Alternatively, an argumentative definition might be simply biased in favor of a particular conclusion in an argument.

- A Libertarian is someone who doesn't understand what a government is.
- A stool is a chair designed to be uncomfortable.
- A laptop is a posture-killing personal computer.

Definition: Neutral definitions

Neutral definitions are at least attempts at trying to define a word or phrase or concept without biasing the reader toward one or another stance toward the thing being defined. They are attempts at bias-free definition (which is probably impossible, but they're at least attempts).

- A drawer is a container on fixed rollers or rails.
- A dog is a domesticated member of the *canus* genus.
- A computer is anything that is designed to perform computations.

We need to be able to ask the right questions of a definition. These three distinctions correspond to good questions we can ask when trying to evaluate a particular definition. Knowing each of these distinctions helps one make better evaluations of definitions and therefore helps one to understand concepts and arguments that rest of definitions more clearly.

Fallacy of Equivocation

From Matthew J. Van Cleave's Introduction to Logic and Critical Thinking, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

Consider the following argument:

Example [Math Processing Error]

Children are a headache. Aspirin will make headaches go away. Therefore, aspirin will make children go away.

This is a silly argument, but it illustrates the fallacy of equivocation. The problem is that the word “headache” is used equivocally—that is, in two different senses. In the first premise, “headache” is used figuratively, whereas in the second premise “headache” is used literally. The argument is only successful if the meaning of “headache” is the same in both premises. But it isn’t and this is what makes this argument an instance of the fallacy of equivocation.

Here’s another example:

Example [Math Processing Error]

Taking a logic class helps you learn how to argue. But there is already too much hostility in the world today, and the fewer arguments the better. Therefore, you shouldn’t take a logic class.

In this example, the word “argue” and “argument” are used equivocally. Hopefully, at this point in the text, you recognize the difference. (If not, go back and reread section 1.1.)

The fallacy of equivocation is not always so easy to spot. Here is a trickier example:

Example [Math Processing Error]

The existence of laws depends on the existence of intelligent beings like humans who create the laws. However, some laws existed before there were any humans (e.g., laws of physics). Therefore, there must be some non-human, intelligent being that created these laws of nature.

The term “law” is used equivocally here. In the first premise it is used to refer to societal laws, such as criminal law; in the second premise it is used to refer to laws of nature. Although we use the term “law” to apply to both cases, they are importantly different. Societal laws, such as the criminal law of a society, are enforced by people and there are punishments for breaking the laws. Natural laws, such as laws of physics, cannot be broken and thus there are no punishments for breaking them. (Does it make sense to scold the electron for not doing what the law says it will do?)

As with every informal fallacy we have examined in this section, equivocation can only be identified by understanding the meanings of the words involved. In fact, the definition of the fallacy of equivocation refers to this very fact: the same word is being used in two different senses (i.e., with two different meanings). So, unlike formal fallacies, identifying the fallacy of equivocation requires that we draw on our understanding of the meaning of words and of our understanding of the world, generally.

The following is from: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

Typical of natural languages is the phenomenon of homonymy24: when words have the same spelling and pronunciation, but different meanings—like ‘bat’ (referring to the nocturnal flying mammal) and ‘bat’ (referring to the thing you hit a baseball with). This kind of natural-language messiness allows for potential fallacious exploitation: a sneaky debater can manipulate the subtleties of meaning to convince people of things that aren’t true—or at least not justified based on what they say. We call this kind of maneuver the fallacy of equivocation

Here’s an example. Consider a banker; let’s call him Fred. Fred is the president of a bank, a real big-shot. He’s married, but he’s not faithful: he’s carrying on an affair with one of the tellers at his bank, Linda. Fred and Linda have a favorite activity:

they take long lunches away from their workplace, having romantic picnics at a beautiful spot they found a short walk away. They lay out their blanket underneath an old, magnificent oak tree, which is situated right next to a river, and enjoy champagne and strawberries while canoodling and watching the boats float by.

One day—let's say it's the anniversary of when they started their affair—Fred and Linda decide to celebrate by skipping out of work entirely, spending the whole day at their favorite picnic spot. (Remember, Fred's the boss, so he can get away with this.) When Fred arrives home that night, his wife is waiting for him. She suspects that something is up: "What are you hiding, Fred? Are you having an affair? I called your office twice, and your secretary said you were 'unavailable' both times. Tell me this: Did you even go to work today?" Fred replies, "Scout's honor, dear. I swear I spent all day at the bank today."

See what he did there? 'Bank' can refer either to a financial institution or the side of a river—a river bank. Fred and Linda's favorite picnic spot is on a river bank, and Fred did indeed spend the whole day at that bank. He's trying to convince his wife he hasn't been cheating on her, and he exploits this little quirk of language to do so. That's equivocation.

A similar linguistic phenomenon can also be exploited to equivocate: polysemy (Greek word, meaning 'many signs (or meanings)'). This is distinct from, but similar to, homonymy. The meanings of homonyms are typically unrelated. In polysemy, the same word or phrase has multiple, related meanings—different senses. Consider the word 'law'. The meaning that comes immediately to mind is the statutory one: "A rule of conduct imposed by authority." (From the *Oxford English Dictionary*) The state law prohibiting murder is an instance of a law in this sense. There is another sense of 'law', however; this is the sense operative when we speak of scientific laws. These are regularities in nature—Newton's law of universal gravitation, for example. These meanings are similar, but distinct: statutes, human laws, are prescriptive; scientific laws are descriptive. Human laws tell us how we ought to behave; scientific laws describe how things actually do, and must, behave. Human laws can be violated: I could murder someone. Scientific laws cannot be violated: if two bodies have mass, they will be attracted to one another by a force directly proportional to the product of their masses and inversely proportional to the square of the distance between them; there's no getting around it.

A common argument for the existence of God relies on equivocation between these two senses of 'law':

[Math Processing Error]

This argument relies on fallaciously equivocating between the two senses of 'law'—human and natural. It's true that human laws are by definition imposed by an authority; but that is not true of natural laws. Additional argument is needed to establish that those must be so imposed.

A famous instance of equivocation of this sort occurred in 1998, when President Bill Clinton denied having an affair with White House intern Monica Lewinsky by declaring forcefully in a press conference: "I did not have sexual relations with that woman—Ms. Lewinsky." The president wanted to convince his audience that nothing sexually inappropriate had happened, even though, as was revealed later, lots of icky sex stuff had been going on. He does this by taking advantage of the polysemy of the phrase 'sexual relations'. In the broadest sense, the phrase connotes sexual activity of any kind—including oral sex (which Bill and Monica engaged in). This is the sense the president wants his audience to have in mind, so that they're convinced by his denial that nothing untoward happened. But a more restrictive sense of 'sexual relations'—a bit more old-fashioned and Southern usage—refers specifically to intercourse (which Bill and Monica did not engage in). It's this sense that the president can fall back on if anyone accuses him of having lied; he can claim that, strictly speaking, he was telling the truth: he and Monica didn't have 'relations' in the intercourse sense. Clinton later admitted to "misleading" the American people—but, importantly, not to lying.

The distinction between lying and misleading is a hard one to draw precisely, but roughly speaking it's the difference between trying to get someone to believe something false by saying something false (lying) and trying to get them to believe something false by saying something true but deceptive (misleading). Besides homonymy and polysemy, yet another common linguistic phenomenon can be exploited to this end. This phenomenon is implicature, identified and named by the philosopher Paul Grice in the 1960s (See his Studies in the Way of Words, 1989, Cambridge: Harvard University Press). Implicatures are contents that we communicate over and above the literal meaning of what we say—aspects of what we mean by our utterances that aren't stated explicitly. People listening to us infer these additional meanings based on the assumption that the speaker is being cooperative, observing some unwritten rules of conversational practice. To use one of Grice's examples, suppose your car has run out of gas on the side of the road, and you stop me as I walk by, explaining your plight, and I say, "There's a gas station right around the corner." Part of what I communicate by my utterance is that the station is open and selling gas right now—that you can go there and solve your problem. You can infer this content based on the assumption that I'm being a cooperative conversational partner; if the station is closed or out of gas—and I knew it—then I would be acting unhelpfully, uncooperatively. Notice, though, that this content is not part of what I literally said: all I told you is that there is a gas station around the corner, which would still be true even if it were closed and/or out of gas.

Implicatures are yet another subtle aspect of meaning in natural language that can be exploited. So a final technique that we might classify under the fallacy of equivocation is false implication—saying things that are strictly speaking true, but which communicate false implicatures. Grocery stores do this all the time. You know those signs posted under, say, cans of soup that say “10 for $10”? That’s the store’s way of telling us that soup’s on sale for a buck a can; that’s right, you don’t need to buy 10 cans to get the deal; if you buy one can, it’s $1; 2 cans are $2, and so on. So why not post a sign saying “$1 per can”? Because the 10-for-$10 sign conveys the false implicature that you need to buy 10 cans in order to get the sale price. The store’s trying to drive up sales.

A striking example of false implicature is featured in one of the most prominent U.S. Supreme Court rulings on perjury law. In the original criminal case, a defendant by the name of Bronston had the following exchange with the prosecuting attorney: “Q. Do you have any bank accounts in Swiss Banks, Mr. Bronston? A. No, sir. Q. Have you ever? A. The company had an account there for about six months, in Zurich.” (Bronston v. United States, 409 US 352 - Supreme Court 1973). As it turns out, Bronston did not have any Swiss bank accounts at the time of the questioning, so his first answer was strictly true. But he did have Swiss bank accounts in the past. However, his second answer does not deny this. All he says is that his company had Swiss bank accounts—an answer that implicates that he himself did not. Based on this exchange, Bronston was convicted of perjury, but the Supreme Court overturned that conviction, pointing out that Bronston had not made any false statements (a requirement of the perjury statute); the falsehood he conveyed was an implicature. (The court didn’t use the term ‘implicature’ in its ruling, but this was the thrust of their argument.)

2.3: Necessary and Sufficient Conditions

It is so easy to confuse Necessary and Sufficient Conditions. Knowing the difference can help you avoid this fate. After introducing the distinction, we'll discuss why it is important to be aware of it.

First, some vocabulary:

A **state-of-affairs** or **event** or **condition** is something that can happen. My going to the mall, or your passing your exam, or the wall's turning red in color, Chris's getting a DUI, an electron's being annihilated, etc. are all examples of states of affairs or events.

When one of these in fact occurs or happens or is true, we say that it **obtains**. So, I can say that

"the state of affairs where Andrew Lavin is an instructor at FRC has ***obtained***." This means that Andrew Lavin is *in fact* an instructor at FRC. 'To **obtain**' is simply the verb that attaches to states of affairs and events and the like. It means something like "happens."

- For Chris getting a DUI to **obtain**, he must have been caught driving while intoxicated.
- The wall turning red will **obtain** if we paint it red.

These are unnatural ways to talk, but this verb can be helpful for defining Necessary and Sufficient Conditions.

Necessary Conditions

A necessary condition is something that **must** obtain if something else *is to* obtain.

It means if that for which it is a necessary condition obtains, then it **must also** obtain/have obtained.

If X is a necessary condition for Y, then any time Y obtains, X must also obtain.

Y ***requires*** or ***needs*** X in order to obtain. (X is necessary for Y)

If peanut butter is necessary for making a PB&J, then you must have peanut butter if you are going to make a PB&J. Any time you make a PB&J, you must have peanut butter. A PB&J requires or needs peanut butter to be a PB&J.

Examples

Example *[Math Processing Error]*

Peanut Butter is a necessary condition for making a PB&J sandwich.

Jelly is also a necessary condition for making a PB&J.

Air being present is a necessary condition for breathing. It's also necessary for a flame.

Paying for a product is a necessary condition for walking out of a store with that product without stealing (and without that product being free or given to you by the store owner).

Proving someone's guilt beyond a reasonable doubt is a necessary condition for convicting them of a criminal offense in the American criminal trial system.

Being an animal is a necessary condition for being a mongoose.

More intuitively, a necessary condition is something you *need* if you are going to do something else. It's whatever *must* happen if something else is to happen. You can't be a mongoose without being an animal, so being an animal is a necessary condition for being a mongoose. If you didn't steal something, then normally it follows that you did the necessary things: you paid for it before walking out of the store.

Sufficient Conditions

A Sufficient Condition is something that is **enough** for something else to obtain.

It means that **if the sufficient condition obtains, then** that for which it is a sufficient condition is sure to obtain as well.

If X is a sufficient condition for Y, then **any time X obtains**, Y will also obtain.

X ***causes*** or **is enough for** Y to obtain. (X is sufficient for Y)

If you break your leg, it will hurt, so breaking your leg is sufficient for being in pain. Breaking your leg is enough (it's all you need to do) to be in pain. If you break your leg (any time you break your leg), you will be in pain. Breaking your leg causes you to be in pain.

Sufficient Conditions are a bit trickier when it comes to examples. They always rely on certain unstated background assumptions and so it feels like most of them have relatively easy-to-find counterexamples. We have to practice a bit of charity in interpreting sufficient conditions.

Examples

Example [Math Processing Error]

Having Peanut Butter, Jelly, Bread, and a knife is a sufficient condition for making a PB&J sandwich (all else being equal).

- Note that these are *individually necessary* and **jointly sufficient**, meaning taken together, they form a sufficient condition.

Taking air in through your mouth into your lungs and extracting the oxygen is a sufficient condition for breathing.

Walking out of a store with one of their products without paying is a sufficient condition for stealing (assuming it wasn't free or given to you by the owner).

Being caught by police in the middle of committing a crime is generally a sufficient condition for being arrested for that crime.

Being a rabbit is a sufficient condition for being a mammal.

If a thing is a rabbit, then it is a mammal. All that is needed to be a mammal is to be a rabbit. If you want to breathe, all that you need to do is take air into your lungs through your mouth and extract the oxygen. That's not the only way to breathe (you can breathe through your nose, for instance), so we know that this is a **sufficient but not necessary** condition.

Necessary and Sufficient Conditions

Note that a condition can be:

1. Necessary *but not Sufficient*

2. Sufficient *but not Necessary*

3. Necessary *and* Sufficient

4. Neither Necessary *nor* Sufficient

Necessary and Sufficient conditions are things that are **both enough for and required for** something else.

If X is a necessary and sufficient condition for Y, then:

> If X obtains, then Y must obtain (so any time X obtains, Y also obtains)
>
> And
>
> If Y obtains, then X must also obtain (so any time Y obtains, X also obtains)

Examples

Example [Math Processing Error]

Putting peanut butter on one slice of bread and putting jelly on another slice of bread and putting them together (PB&J sides inwards) is a necessary and sufficient condition for making a PB&J sandwich.

Taking sandwich ingredients and sandwiching them between two slices of bread or two halves of a roll is a necessary and sufficient condition for making a sandwich.

Having a belief that is true and isn't accidentally true is a necessary and sufficient condition for having knowledge (some think).

Earning a final grade of C or better is a necessary and sufficient condition for passing a General Education course.

Being 21 years of age or older is a necessary and sufficient condition for being legally allowed to drink in the US (barring special legal constraints and not counting states that allow certain cases of underage drinking).

Neither

Of course, there are loads of conditions that are neither necessary nor sufficient for other things to occur. Here are some examples

Example [Math Processing Error]

Claude pouring his espresso too fast is neither necessary nor sufficient for the sun to be in the sky.

Eating BBQ is neither necessary nor sufficient for being vegetarian (in fact, depending on the protein, it's probably sufficient for *not* being a vegetarian!)

Paying for a milkshake after ordering it for yourself is neither necessary nor sufficient to get a ride to Phoenix with your cousin for Christmas.

Background Assumptions

Every conditional claim makes certain *background assumptions*. For instance, driving is a sufficient condition to get to work on time. But it's not going to be sufficient if your car breaks down, or if you get in an accident, or if a meteor destroys your workplace, or if Thanos from Marvel's Avengers turns you into dust while you're on the middle of the drive. Every time we claim something is a condition for something else—especially a sufficient condition—we are making the assumption that everything else basically proceeds as normal. If weird stuff starts happening, then my condition might not be sufficient anymore—but this doesn't mean that what I said originally was false.

Who Cares?

Why study Necessary and Sufficient conditions?

First, it helps us understand the dependencies between events and things in our world. Getting clearer about these distinctions helps us to reason more clearly about conditions, conditionals (if...then... propositions), cause and effect, evidence and implication.

Second, it helps us get clear on what is involved in definition. What are we trying to do in defining something? Generally, we're trying to give necessary and sufficient conditions for that thing (though this is a problematic account of definition, it's good enough for our purposes). When we define what a dog is, we're listing what all dogs must have to be dogs (dog DNA, Dog parents, etc.) and what features are such that if anything has those features, then that thing is a dog (what features make something a dog). It might be necessary and sufficient for being a dog that one has a certain canine genetic signature. All and only dogs have this signature. This is almost certainly false, since biology is a wild and confusing place, but we get the idea of what we're at least trying to do when we define a term: identify what all and only those things that fall under that concept have.

Third, it helps us *avoid common fallacies* (errors in reasoning). Confusing necessary conditions for sufficient conditions is all-too-common in everyday reasoning and (as you'll see in the videos from Wireless Philosophy) can lead to really bad consequences—like wrongful convictions!

Here's an example:

Example [Math Processing Error]

If we find the murder weapon in Scott's apartment, then we know Scott committed the crime.

Is finding the murder weapon necessary or sufficient? What happens if we don't find the murder weapon? Is Scott off the hook? If we know that either Scott or Mohinder committed the crime, but we don't find the murder weapon in Scott's apartment, does that mean Mohinder did it?

If we don't get clear on the relationship between necessary and sufficient conditions, then we won't be able to answer these questions properly and we'll potentially convict the wrong person or let the wrong person go.

In this case, finding the murder weapon is a sufficient condition for knowing that Scott committed the crime.

If we don't find it, that doesn't mean anything at all when it comes to whether or not Scott committed the crime.

Let's think about an analogy: If I poison you, you'll get sick. If I don't poison you, then who knows if you'll get sick? We don't know that you won't get sick! You might eat bad mayonnaise or get an infected tick bite!

So, we don't know that Scott didn't commit the crime if we didn't find the murder weapon in his apartment.

Here are some other examples of necessary and sufficient condition confusion:[1]

Example *[Math Processing Error]*

Juan: "How do you think you'll do on our philosophy exam tomorrow?"

Monique: "Great, I read all the books."

Juan: "Yeah but do you understand this stuff?"

Monique: "I said I read all the books, didn't I?"

Monique is confusing the fact that reading all of the books is *necessary* for understanding the material of the course with the false idea that reading all of the books is *sufficient.* It's not sufficient, you have to read *and* understand (and probably some other stuff too!) in order to do well on the philosophy exam.

Example *[Math Processing Error]*

Don't let all the talk about the necessity of exercise to a long life mislead you. Jim was a jock all his very short life.

Jim dies tragically short, but not from exercise! The speaker here confuses the *fact* that regular exercise is *necessary* for a long, healthy life with the false claim that regular exercise is *sufficient*. It's not sufficient, as Jim's tragic example shows. It takes a lot more than regular exercise to get you a long and healthy life since failing to exercise isn't the only thing that can go wrong.

Example *[Math Processing Error]*

I don't know why the car won't run; I just filled the gas tank.

Again, a full gas tank is *necessary*, but it isn't *sufficient* for a car to run. There is a lot more that can go wrong with a car other than an empty gas tank. Having gas isn't the only condition that needs to be in place.

Example *[Math Processing Error]*

She didn't have any diseases, I don't understand why she died!

Having a disease is sometimes *sufficient* for dying (tragic, though it may be). But it isn't *necessary*. There are many more ways to die (sadly) than getting a disease.

Example *[Math Processing Error]*

My cat didn't get taken away by an eagle, so I don't understand why it's missing.

This is a kind of comical example. It's silly, right? Of course, if your cat did get taken away by an eagle, then it would be missing, but it would also be missing if it was hiding under the porch, or hitched a ride on a semi-truck to Tucson! There are lots of ways for a cat to go missing (so it's not necessary that it be taken away by an Eagle), but it isn't indeed *sufficient* for it to be missing that an Eagle swooped in and grabbed it.

---

[1] I got most of these from <http://www.txstate.edu/philosophy/resources/fallacy-definitions/Confusion-of-Necessary.html>

2.4: Chapter 2 - Key Terms

- **Vagueness**
- **Ambiguity**
- **Intension**
- **Extension**
- **Synthesis**
- **Analysis**
- **Stipulative vs Descriptive definitions**
- **Too broad or narrow vs. apt definitions**
- **Argumentative vs Neutral Definitions**
- **Necessary Conditions**
- **Sufficient Condition**

2.E: Chapter Two (Exercises)

Exercise *[Math Processing Error]*: Vagueness vs Ambiguity

Is each inference an illustration of vagueness or ambiguity?

A. She said "go blow your nose" and I've been trying for an hour to blow on my nose. It's tricky.

B. I don't think I lied, I said "I'll be there to help you if you need it" and you didn't *really need* help this time since you clearly survived without my help.

C. Teacher, you said to do the test "without help" and Jesus is using a pencil to complete the test. That's helping him do it!

D. You said not to tattle, so when I saw the murderers running away from their victim I remained tight-lipped as ever.

E. Uggh. I said I wanted regular blonde hair, not platinum blonde hair! You messed my hair up.

F. The physics teacher said that on a level plane it's indeterminate which way a stationary object will start to roll, but I've been setting this ball down carefully on the floor of this airplane and it has been rolling backwards every time!

G. Of course humans can fly. I just flew to Los Angeles last week!

H. She said "I'll meet you inside," so I went deep inside my soul in search of her, but never found her.

I. "As I was leaving this morning, I said to myself, 'The last thing you must do is forget your speech.' And, sure enough, as I left the house this morning, the last thing I did was to forget my speech."– Rowan Atkinson

J. You said we'd meet early today, but I've been waiting here since 5 am and you just now showed up at 7 am!

K. "I didn't go to the party because Shamik was there."

"Oh, so you went to sample the appetizers?"

L. I'm not a very good student, I only have a 3.87 GPA.

M. You're so rich! You have a working refrigerator.

N. It's so cold outside today, it's like 50 degrees.

O. "Call me a taxi, please."

"Okay, you're a taxi!"

P. Union demands increased unemployment. -newspaper headline

Q. My mom says I don't listen to her. Well I do listen to her, I just don't do what she tells me to do!

Exercise *[Math Processing Error]*: Intensional Meaning and Extensional Meaning

Is each of the following series, determine whether the extension is increasing or decreasing and then determine whether it has more intensional information (it's getting more specific) or decreasing intensional information (it's getting less specific). "No change" is also an option for both dimensions (intension and extension).

A. Indian food, North Indian food, Masala dishes, Paneer Masala.

B. My rude neighbor guy, rude guys, rude people, people.

C. The Moon, moons, moons in our solar system, natural satellites (any natural object orbiting a planet) in our solar system, natural satellites.

D. Jimmy Carr (an English comedian), English comedians, British comedians, comedians.

E. Earth, planets in our solar system with intelligent life, the planet on which Captain America: Civil War takes place, the planet that is home to almost all living human beings.

F. Pink Floyd, rock groups from the 70's and 80's, English-speaking rock groups, rock groups

? Exercise [Math Processing Error]: Types of Definitions

For each definition, determine the answer to each of the following questions: a) Is it Stipulative or Descriptive? b) Is it too broad or too narrow, or is it apt? c) Is it argumentative or Neutral?

A. An immigrant is any person who comes into a country for a stay that is longer than two months without being a citizen of the country.

B. An immigrant is someone like Natalie Portman, Gal Gadot, John Oliver, or James McAvoy.

C. Furniture is a human-made object used to store or hold things; or sit or lay on.

D. A ruler is a straight and rigid device used to measure lengths.

E. A criminal is someone who has committed a crime.

F. A Russian is someone who is not to be trusted.

G. A communist is someone like Stalin, Pol Pot, Mao, or Kim Jong Un.

H. A wristwatch is a time piece made to be worn around one's wrist.

I. A smart device is any device with the ability to connect to the internet.

? Exercise [Math Processing Error]: Fallacy of Equivocation

Explain how this is equivocation...

A. You said you would meet me at the park, but we've already met—years ago in fact—so I don't think we should plan to meet in the future.

B. That bumper sticker says "the cost of freedom is great," which is weird because I never think of war as all that great. It's pretty awful.

C. Peter: Let's go to the gym to do a bit of working out.

Tamil: Okay, but you won't need your running shoes at the climbing gym, you should grab your climbing shoes instead and just wear street shoes on the way there.

? Exercise [Math Processing Error]: Necessary and Sufficient Conditions

Fill in each blank with either Necessary and Sufficient, Necessary but not Sufficient, Sufficient but not Necessary, or Neither Necessary nor Sufficient.

A. Giving birth a baby is a _______ condition to becoming a mother.

B. Eating spicy foods is, for many people, a _______ condition for having gastro-intestinal problems.

C. Drinking milk is, for lactose-intolerant people, a ________ condition for having gastro-intestinal problems.

D. Flying in a plane is a ________ condition for getting from California to Hawaii in under 24 hours.

E. Dropping a sick beat is a ___________ condition for rapping.

F. Being alive is a ________ condition for running for President.

G. Eating at least every three weeks is a ___________ condition for staying alive (for humans).

H. Having an ear of corn in your possession is a _______________ condition for making corn bread.

I. Selling drugs in front of a police officer is, in normal circumstances, a ___________ condition for being arrested

J. Being a set of bound pages containing a more or less unified piece of writing or collection of writings is a _____________ condition for being a book.

K. Putting fruit, yogurt, and milk in a blender and blending until smooth is a ________ condition for making a smoothie.

CHAPTER OVERVIEW

3: Argument Mapping

3.1: The Basics

So far, we've discussed the basic ideas behind arguments or inferences. Each argument has ***premises*** which are the *assumptions* or the *support* of the argument. Each argument also has usually one, but sometimes more ***conclusions***. The conclusion is the *main point* of the argument. The goal of any argument is to offer reasons for believing the conclusion. The reasons are the premises and the claim that you are supposed to accept if you agree with the argument is the conclusion.

So far so good. But there's a lot more that we can say about arguments.

Ideally, when we're trying to understand an argument fully—long before we decide whether or not we *agree* with the argument or whether or not it's a *good* argument—we have a full grasp of the **structure** of the argument. That is, we need to know which premises go with which other premises, whether each premise is supposed to directly demonstrate the conclusion or is merely indirect support for the conclusion, etc. In short, we need a **map** or a **diagram** of the argument *before* we can decide whether or not it's a good argument.

Simple arguments are called syllogisms: 2 premises and 1 conclusion and immediate inferences: 1 premise and 1 conclusion.

Like:

I like all vegetables

Carrots are a vegetable

So I like Carrots

Or:

I like all vegetables

So, there aren't any vegetables I don't like.

But normal arguments (arguments you'd find in a letter to the editor or in a social media post or on the radio or tv) aren't like that—they have more premises, some of which don't directly support the conclusion, but instead support other premises. It's like a big complex argument that's actually made out of smaller arguments.

So, if you want to understand how a complex argument in the real world hangs together, you need to be able to construct a map or diagram of that argument.

We'll need to find out two things about each premise:

1. **What kind of support does it offer for its conclusion?** Does it support its conclusion in conjunction with other premises? Or does it instead form an argument by itself for the conclusion?

2. **Does it support the main conclusion directly?** Or does it instead support the conclusion *indirectly* by offering support for another premise, which in turn supports the main conclusion?

How do we go about actually building an Argument Map? Well, we could choose any convention at all, so we have to decide on what sorts of shapes, labels, symbols, etc. we'll use for the sake of this course.

The first thing to note is that some people teach argument mapping going in an upwards direction—meaning that the conclusion would be on top and the premises for the conclusion would be below it. But we're going to go a different way so that our argument maps more clearly track the usual format of an argument: the premises on top and the conclusion on bottom.

Here are some basic concepts and the associated conventional symbols and shapes:

We use arrows to signify **Inferential Links** or support. Every arrow means "**implies that**" or "**therefore**". Read backwards (upwards), this diagram to the right means: "1 is true" "why?" "because of 2". Or "Given that 2 is true, 1 follows."	Paradigm Example: (1) We can go now, because (2) the car is packed.	

Paradigm Example: (1) I know that Voodoo is real, because (2) My cousin saw someone take on the characteristics, personality, and voice of a spirit during a ceremony. (3) My cousin told me that she saw this last week.		Sometimes, we might find that a premise offers indirect support for the main conclusion of the argument. In that case, we have to build a vertical pattern into our argument map that might look something like the diagram to the left.

Conjoint vs. Independent Support

We need to be able to decide (once we've sorted out which are premises for which conclusion) what kind of support a set of premises provide for their conclusion. **It's independent support when each premise seems like it's an argument for the conclusion on its own.** *It's conjoint support when a premise doesn't seem to support the conclusion without the help of the other premises.* A good test for conjoint support is to pretend one of the premises is *false*. Does this affect the inference(s) from the other premise(s) to the conclusion?

Labradors are gentle, but they aren't very aggressive,

so they wouldn't make good guard dogs.

This feels like independent support because each inference makes sense on its own:

Labradors are gentle,

so they wouldn't make good guard dogs.

Labradors aren't very aggressive,

so they wouldn't make good guard dogs.

Let's look at another:

[1] Vegetables are healthy and [2] tomatoes are vegetables, so [3] tomatoes are healthy.

Since 1 is a general principle and 2 is an instance of that general principle (or something like that), it makes sense to think that they're conjoint. Any time you see this pattern—where one premise is a definition or general claim and another premise is a more particular claim that falls under that definition or general claim—you'll think that those premises are likely conjoint.

The "General-Specific Pattern

When you see two premises where one premise is a general definition, a generalization, a hypothetical or conditional, or a general principle, and the other premise is a specific claim about an individual under that generalization, those are almost certain to be conjoint premises.

Examples

- A motorbike is any two-wheeled motor-driven vehicle and that moped has two wheels that are driven by a motor, so...
- If anyone goes to the amusement park, they're going to be exhausted at the end of the day; and Cheri went to Six Flags today, so...
- Lying is wrong, but getting out of trouble would require me to lie, so....

If we try negating 2, then the inference doesn't make any sense:

[1] Vegetables are healthy and ***[2] tomatoes are not vegetables****, so [3] tomatoes are healthy.*

What????

If we try negating 1, the inference falls apart again:

[1] Vegetables are unhealthy *and [2] tomatoes are vegetables, so [3] tomatoes are healthy.*

What??????

Let's try one more slightly more complex conjoint support example:

Example [Math Processing Error]

[1] Gina told me the Earth is round and [2] Gina wouldn't lie to me, and furthermore [3] Gina is an astrophysicist, so [4] the Earth is round.

Let's try the negation test on 1:

***[1] Gina told me the Earth is flat** and [2] Gina wouldn't lie to me, and furthermore [3] Gina is an astrophysicist, so [4] the Earth is round.*

What??? Let's try it on 2:

*[1] Gina told me the Earth is round and **[2] Gina often lies to me**, and furthermore [3] Gina is an astrophysicist, so [4] the Earth is round.*

What???? Let's try it on 3:

*[1] Gina told me the Earth is round and [2] Gina wouldn't lie to me, and furthermore **[3] Gina is not an astrophysicist**, so [4] the Earth is round.*

Well... this isn't as incoherent as the other examples. But why mention that Gina is an astrophysicist at all if it doesn't at least *help* 1 and 2 demonstrate the conclusion that the Earth is round? With the negation of 3 as part of the argument, it seems thoroughly awkward that we should be talking about Gina being or not being an astrophysicist at all. If anything, it seems to *work against* the inference.

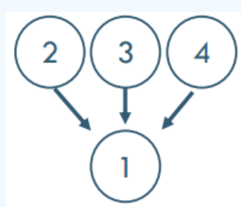
What's the lesson here? The negation test isn't perfect, but it does almost always reveal when you've got a premise that seems to work together with other premises. In the Gina case, we've got a premise that is **closely related in subject matter** and so we've got some reason to conjoin it with 1 and 2.

Here's how we go about mapping conjoint vs. independent support once we've decided what sort of support is involved.

Mapping Independent Support

Example [Math Processing Error]

We use **multiple arrows** to signify multiple **independent** inferences. So, we have many premises which *do not work together to demonstrate the conclusion*. Each premise offers its own reason for accepting the conclusion.



Paradigm example:

(1) This test is easy.

(2) Tetsuo got an A on the test and

(3) Xochitl got an A on the test and

(4) Francisco got an A on the test.

If the other premises were not there, the argument would not fall apart. The premises don't need each other to be true to support the conclusion.

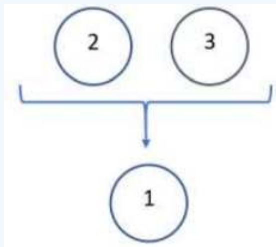
"Given 2, 1 follows, and given 3, 1 follows, and given 4, 1 follow."

Independent support is really like having multiple inferences. So the map above seems to tell us that there are three separate inferences that just happen to have the same conclusion.

Mapping Conjoint Support

Example [Math Processing Error]

We use **brackets** to signify a single inference with many **conjoint or mutually dependent premises**. *The premises work together to support the conclusion*. Without the other conjoint premises, it would be unclear why one conjoint premise should be taken as a reason for accepting the conclusion.



Paradigm example:

(1) You are behaving unfairly.

(2) You're giving more to some than to others and

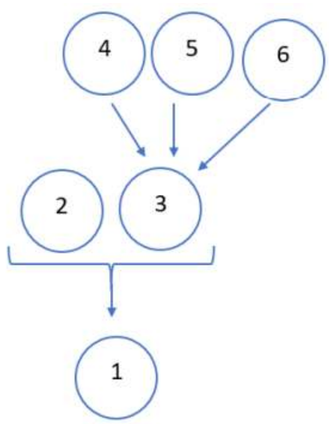
(3) giving more to some than to others isn't fair.

If any one of them is false or wasn't there to begin with, the inference falls apart.

"Given 2, 1 doesn't follow unless we also have 3 (and 4, 5, 6, ...)."

Deductive arguments are more often than not conjoint support. This is just a rough and ready rule, but the way standard Deductive arguments (without extra irrelevant premises) work is that the premises are all necessary for the inference to demonstrate the conclusion. So it makes sense that they would be conjoint premises.

Here's a complete example of a problem like you might see on a quiz or exam (though they'll usually be less complex than these, at least to start out).



Example [Math Processing Error]

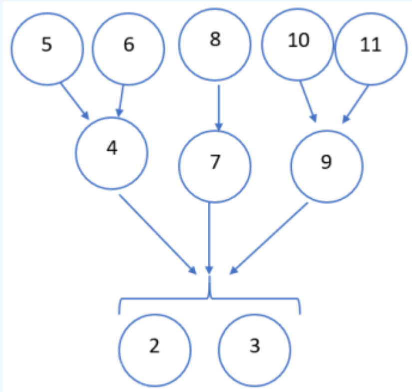
(1) Government mandates for zero-emission vehicles won't work because (2) only electric cars qualify as zero-emission vehicles, and (3) electric cars won't sell. (4) They are too expensive, (5) their range of operation is too limited, and (6) recharging facilities are not generally available.

Adding in 3 makes the inference make sense again (*Oh, I see, electric cars won't solve our problems*). You can do the same by taking 2 away. Wait, we'll say, what about *other zero-emission vehicles*??? Adding 2 back in makes sense of the inference.

4, 5, and 6 are independent because they don't have much to do with one another. The inference from 4 to 3, 5 to 3, and 6 to 3 all makes sense. "They're too expensive, so they won't sell." (makes sense). "Their range is limited, so they won't sell" (makes sense). "There aren't enough recharging facilities, so they won't sell" (makes sense!).

Example [Math Processing Error]

We also use **downward braces** if there are more than one conclusion for any given inference. This is called **Multiple Conclusions**.



Example:

(1) The president may have her faults, but

(2) she is an outstanding leader and

(3) we should reelect her.

(4) Her foreign policy has brought about respite from violence in various war torn regions as

(5) she sent in troops to protect refugees in Rwanda and (6) she negotiated an armistice between Egypt and Israel. (7) Her economic policy has also been largely successful in that (8) a potential recession has been avoided for now. (9) She is also a great moral leader as (10) hers is a model family and (11) she demonstrates true integrity daily.

Notice how 1 isn't actually part of the argument: it just introduces the topic but isn't a premise or conclusion. 2 and 3 are both conclusions (notice the "and", which often links premises to premises and conclusions to conclusion) because neither is a premise/evidence for the other and both are implied by the rest of the argument (4, 7, and 9).

Why did we go with independent support for all of the top-most premises? Try to reason through it on your own.

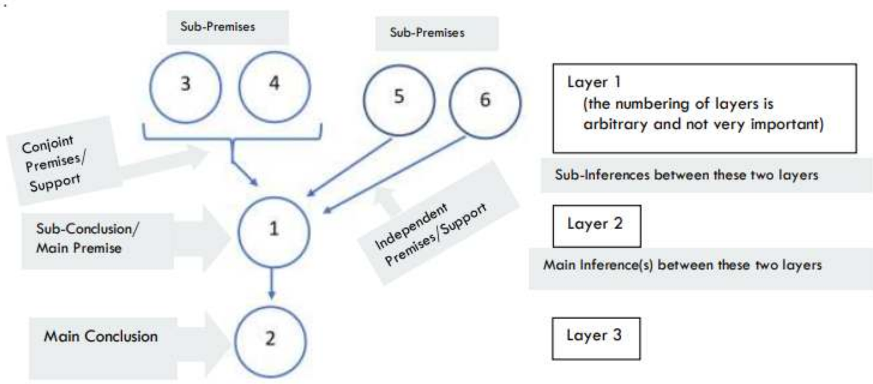
Terminology

Let's introduce some new terminology so we can have a common language with which to talk about arguments:

- A "**level**" or "**layer**" of an argument map is one horizontal row of a carefully-drawn argument map. Notice how the previous argument map above is drawn so that even though there's a lot going on in the argument, we can see 3 distinct layers or horizontal rows?
- A **Main Conclusion** is the final conclusion of the argument. It doesn't serve as a premise/support for any other proposition in the complex argument. It's always in the bottom-most layer
- A **Main Premise** is one among the set of premises that directly support the main conclusion. It's always in the layer that's the second from the bottom.
- A **Sub-Inference** is an inference from a premise to another premise. The conclusion of a sub-inference is never in the bottom-most layer.
  - A **sub-premise** is a premise in a sub-inference.

- A **sub-conclusion** is a conclusion in a sub-inference. (Note that a sub-conclusion is *always* a premise itself, and that it is usually one of the main premises unless the argument gets really complex).

So here it is, the anatomy of a typical 3-layer argument diagram:



The following excerpt from Knachel's text covers some of the same ground we just covered, but sometimes it's helpful to see a different explanation of the same thing:

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

V. Diagramming Arguments

Before we get down to the business of evaluating arguments—of judging them valid or invalid, strong or weak—we still need to do some preliminary work. We need to develop our analytical skills to gain a deeper understanding of how arguments are constructed, how they hang together. So far, we've said that the premises are there to support the conclusion. But we've done very little in the way of analyzing the structure of arguments: we've just separated the premises from the conclusion. We know that the premises are supposed to support the conclusion. What we haven't explored is the question of just how the premises in a given argument do that job—how they work together to support the conclusion, what kinds of relationships they have with one another. This is a deeper level of analysis than merely distinguishing the premises from the conclusion; it will require a mode of presentation more elaborate than a list of propositions with the bottom one separated from the others by a horizontal line. To display our understanding of the relationships among premises supporting the conclusion, we are going to depict them: we are going to draw diagrams of arguments.

Here's how the diagrams will work. They will consist of three elements: (1) circles with numbers inside them—each of the propositions in the argument we're diagramming will be assigned a number, so these circled numbers in the diagram will represent the propositions; (2) arrows pointed at circled numbers—these will represent relationships of support, where one or more propositions provide a reason for believing the one pointed to; and (3) horizontal brackets—propositions connected by these will be interdependent (in a sense to be specified below).

Our diagrams will always feature the circled number corresponding to the conclusion at the bottom. The premises will be above, with brackets and arrows indicating how they collectively support the conclusion and how they're related to one another. There are a number of different relationships that premises can have to one another. We will learn how to draw diagrams of arguments by considering them in turn.

Independent Premises

Often, different premises will support a conclusion—or another premise—individually, without help from any others. When this is the case, we draw an arrow from the circled number representing that premise to the circled number representing the proposition it supports.

Consider this simple argument:

[Math Processing Error] Marijuana is less addictive than alcohol. In addition, [Math Processing Error] it can be used as a medicine to treat a variety of conditions. Therefore, [Math Processing Error] marijuana should be legal.

The last proposition is clearly the conclusion (the word 'therefore' is a big clue), and the first two propositions are the premises supporting it. They support the conclusion independently. The mark of independence is this: each of the premises would still provide support for the conclusion even if the other weren't true; each, on its own, gives you a reason for believing the conclusion. In this case, then, we diagram the argument as follows:

Intermediate Premises

Some premises support their conclusions more directly than others. Premises provide more indirect support for a conclusion by providing a reason to believe another premise that supports the conclusion more directly. That is, some premises are intermediate between the conclusion and other premises.

Consider this simple argument:

[Math Processing Error] Automatic weapons should be illegal. [Math Processing Error] They can be used to kill large numbers of people in a short amount of time. This is because [Math Processing Error] all you have to do is hold down the trigger and bullets come flying out in rapid succession.

The conclusion of this argument is the first proposition, so the premises are propositions 2 and 3. Notice, though, that there's a relationship between those two claims. The third sentence starts with the phrase 'This is because', indicating that it provides a reason for another claim. The other claim is proposition 2; 'This' refers to the claim that automatic weapons can kill large numbers of people quickly. Why should I believe that they can do that? Because all one has to do is hold down the trigger to release lots of bullets really fast. Proposition 2 provides immediate support for the conclusion (automatic weapons can kill lots of people really quickly, so we should make them illegal); proposition 3 supports the conclusion more indirectly, by giving support to proposition 2. Here is how we diagram in this case:

Joint Premises

Sometimes premises need each other: the job of supporting another proposition can't be done by each on its own; they can only provide support together, jointly. Far from being independent, such premises are interdependent. In this situation, on our diagrams, we join together the interdependent premises with a bracket underneath their circled numbers.

There are a number of different ways in which premises can provide joint support. Sometimes, premises just fit together like a hand in a glove; or, switching metaphors, one premise is like the key that fits into the other to unlock the proposition they jointly support. An example can make this clear:

[Math Processing Error] The chef has decided that either salmon or chicken will be tonight's special. [Math Processing Error] Salmon won't be the special. Therefore, [Math Processing Error] the special will be chicken.

Neither premise 1 nor premise 2 can support the conclusion on its own. A useful rule of thumb for checking whether one proposition can support another is this: read the first proposition, then say the word 'therefore', then read the second proposition; if it doesn't make any sense, then you can't draw an arrow from the one to the other. Let's try it here: "The chef has decided that either salmon or chicken will be tonight's special; therefore, the special will be chicken." That doesn't make any sense. What happened to salmon? Proposition 1 can't support the conclusion on its own. Neither can the second: "Salmon won't be the special; therefore, the special will be chicken." Again, that makes no sense. Why chicken? What about steak, or lobster? The second proposition can't support the conclusion on its own, either; it needs help from the first proposition, which tells us that if it's not salmon, it's chicken. Propositions 1 and 2 need each other; they support the conclusion jointly. This is how we diagram the argument:

The same diagram would depict the following argument:

[Math Processing Error] John Le Carre gives us realistic, three-dimensional characters and complex, interesting plots. *[Math Processing Error]* Ian Fleming, on the other hand, presents an unrealistically glamorous picture of international espionage, and his plotting isn't what you'd call immersive. *[Math Processing Error]* Le Carre is a better author of spy novels than Fleming.

In this example, the premises work jointly in a different way than in the previous example. Rather than fitting together hand-in-glove, these premises each give us half of what we need to arrive at the conclusion. The conclusion is a comparison between two authors. Each of the premises makes claims about one of the two authors. Neither one, on its own, can support the comparison, because the comparison is a claim about both of them. The premises can only support the conclusion together. We would diagram this argument the same way as the last one.

Another common pattern for joint premises is when general propositions need help to provide support for particular propositions. Consider the following argument:

We shouldn't elect someone who has proven an *incompetent business leader.*

Candidate Z has proven an incompetent CEO. So, we shouldn't elect Candidate Z.

These premises will be mapped with conjoint support since the premises need to work together to show the conclusion. One general principle about who we shouldn't elect, and one particular claim about Candidate Z.

End Knachel Text

Examples

Let's walk through a few examples of arguments that need mapping:

Example *[Math Processing Error]*

She's the best girlfriend ever. She bought me a new backpack for Christmas, she's never late for a date, and she always treats me with care.

Solution

First, we need to identify each *proposition*—that is, each claim that can be true or false independently of the other claims. This is a bit interpretive, so sometimes there aren't hard and fast rules that produce one particular right answer, but generally we can all come up with the same propositions:

(1) She's the best girlfriend ever. (2) She bought me a new backpack for Christmas, (3) she's never late for a date, and (4) she always treats me with care.

What a nice young person! Next, we need to decide what the conclusion is and which propositions are premises. A nice test that often helps is to read all of the premises and then say "therefore..." and then read what you think is the conclusion. It should make sense as an inference if you do this properly. For instance, this is clearly not so good:

She's the best girlfriend ever, she bought me a new backpack, and she always treats me with care, ***therefore*** *she's never late for a date.*

Uhhhhh...what?

This one sounds a lot more sensical:

She bought me a new backpack, she's never late for a date, and she always treats me with care, ***therefore*** *she's the best girlfriend ever.*

It seems like the three premises are *evidence* for the claim that she is the best girlfriend ever. The thing we're being asked to believe as a result of this reasoning is that she's the best girlfriend ever. So that is the conclusion of the inference.

Now we've already basically ruled out that 2, 3, and 4 have any inferential relationship between them. They all seem to give us reasons for believing the conclusion directly. Furthermore, none of them seems to give us reason for believing any other. *Maybe* 4 could be the conclusion of 2, but that's a real stretch. So based on all of this, we can reasonably conclude that 2, 3, and 4 are all on the same level and are all main premises for the conclusion.

Next, we need to decide if these are conjoint or independent premises. What do you think?

How do we decide? Using the negation test. If negating or saying the opposite of one premise doesn't make the inference fall apart, then the premises are *not conjoint*—they're independent. Let's try it here:

*She bought me a new backpack, **she's sometimes late for a date**, and she always treats me with care, **therefore** she's the best girlfriend ever.*

I mean, it is a bit weird, but it's not *nonsense*. Sure, she's sometimes late for a date, but the inference still makes sense.

***She hasn't bought me a new backpack**, but she's never late for a date, and she always treats me with care, **therefore** she's the best girlfriend ever.*

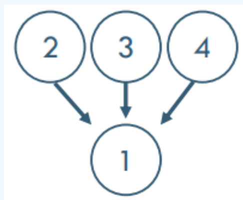
Again, it's strange, but not nonsensical. We wonder why the backpack thing is brought up in the first place, but we don't immediately think "oh, well, she can't be the best girlfriend ever if she hasn't bought you a backpack!" Instead, we just think, "she's clearly an excellent partner, backpack or none."

The last one is a bit stranger:

*She bought me a new backpack, she's never late for a date, **but she doesn't always treat me with care**, **therefore** she's the best girlfriend ever.*

Interesting...the case is definitely pretty weak for her being the bester girlfriend ever at this point, but the inference hasn't utterly fallen apart. An opposite conclusion doesn't now follow, we just have weaker reason for accepting the conclusion than we had before. This test reveals how strong a piece of evidence proposition 4 was for the conclusion in the original argument, but it doesn't tell us that 4 is conjoint—the argument didn't fall apart.

With all of this in mind, the premises appear to be independent reasons from one another for accepting the conclusion that she is the best girlfriend ever. So the argument map looks like so:



How about another example? This time I've skipped right to numbered propositions:

Example [Math Processing Error]

(1) Obama was the best President in American history. (2) He protected people with pre-existing medical conditions from certain financial ruin or death by passing the Affordable Care Act, and (3) that feat was among the greatest legislative victories an American President has ever known. (4) He was able to topple the head of Al-Qaida and the mastermind of the 9/11 attacks, and (5) he oversaw the recovery from the largest economic disaster since the Great Depression. (6) Anyone who could bring us back from the brink of global economic meltdown to a stable and healthy economic like we had at the end of his tenure must be a truly great president.

Solution

Before we ever get to the question of whether or not this is a good argument, or what's wrong with it if anything, or whether or not the conclusion is true, we must *understand* the argument. In particular we must understand the *structure* of the argument. This argument is complex, so what's going on here?

What's the conclusion? It's probably somewhat obvious here. There's one claim that seems like the kind of claim someone might have as a thesis statement, or might defend in an Oxford-style debate. There's one claim that seems to unify the rest of the propositions: everything is meant to justify or defend the claim that Obama was the best President in American history.

With a longer argument like this, sometimes it is best to simply work sentence-by-sentence. 2 and 3 are part of the same sentence. The "and" tells us that there probably is no inferential link between 2 and 3. "and" is usually not interchangeable with "therefore". When we read the content of 2 and 3, furthermore, 3 makes reference to 2. Often when a premise makes reference to another premise we can conclude that they are conjoint premises. Not always, mind you, and often that means that one is a subpremise for the other. Nevertheless, in this case the reference to "that feat" in 3 ties 3 to 2 conjointly. We can run the negative test to be sure we're correct here:

(1) Obama was the best President in American history. (2) He protected people with pre-existing medical conditions from certain financial ruin or death by passing the Affordable Care Act, and (3) that feat was ***an unremarkable legislative accomplishment.***

Now I'm unclear why we should think he's the best president in history if the reason we're being given is that he passed an important, but unremarkable piece of legislation. Not convincing. If anything, it seems to suggest that he was a fine, but unremarkable president.

(1) Obama was the best President in American history. (2) ***He didn't protect people with pre-existing medical conditions from certain financial ruin or death by passing the Affordable Care Act,*** *and (3) that feat* ***would have been*** *among the greatest legislative victories an American President has ever known.*

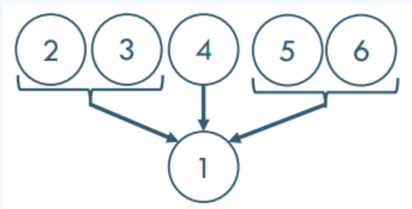
Ummm...no. His *not* passing landmark legislation doesn't make him the best president.

> This is one way you know you're dealing with conjoint premises: if one premise explains how the other premise supports the conclusion.

So these two premises are conjoint. What about 4? It's part of the same sentence as 5, but the topics are so wildly different that it's hard to see how they could be conjoint premises. Instead, it seems safe to assume they're independent and that they're independent from 2 and 3 for the same reason. They do, however, appear to be premises for the main conclusion (1) and so appear to belong on the second level with the other main premises 2 and 3.

The last proposition, though, seems to essentially be about the same topic as 5 and furthermore seems to be the *reason* 5 supports the conclusion. This is one way you know you're dealing with conjoint premises: if one premise explains how the other premise supports the conclusion. So 6 and 5 appear to be conjoint. If you ran the negative test, you'd soon learn that the negated inferences make no sense.

As a result, the whole argument map, which is a bit strange looking, looks like this:



3.2: Missing Assumptions

We need to be able to identify (and then to incorporate into our argument maps) assumptions that are part of the argument, but that weren't explicitly stated. These are called **Hidden Premises**, **Missing Assumptions**, **Suppressed Premises/Assumptions**, etc. Regardless of the name, these are cases where an argument in fact relies on a claim that it doesn't state as a premise. There is a claim that *must* be true if the inference is to make sense, but isn't explicitly claimed to be true by the argument as it is written or spoke.

Identifying Hidden Assumptions

It's a bit tricky, but it might be one of the most important and practical skills you'll learn in this class. How do we figure out *when* an argument has one or more *hidden premises* and how do we identify *what* those premises are?

Well, one answer is simply that formal logic does the job for us when we're dealing with deductive arguments. This is the math-like system of argument analysis we'll learn later in the semester. But for now, we need some informal tools to allow us to identify hidden premises or assumptions. This is not only a skill you can build without knowing logic, but is also a skill that extends to *inductive* reasoning as well and so is far more broadly applicable.

Check out this argument:

Flowers smell nice

???

[Math Processing Error] *Let's plant some flowers*

[that little triangle of dots is a "therefore" sign]

We can't quite get from "smelling nice" to "things we'll plant" without an assumption which *links* these two ideas. Notice how flowers are in both the premise and the conclusion, so we don't need to link the topic of flowers together with the topic of flowers.

Abstractly, is A is related to B, and C is related to B, then what we need is something linking A and C so that we can bridge the gap between A and B being associated with one another to C and B being associated with one another.

Less abstractly, if we have three topics: flowers, smelling nice, and things to plant; then we need something linking smelling nice and things to plant so that we know the fact that flowers smell nice is a compelling reason to think that flowers are the kind of thing we should be planting. We know that flowers smell nice and we're trying to get to flowers should be planted. Any idea what the hidden link might be?

The hidden assumption is something like "we should plant things that smell nice." Can you see how that completes the inference? Check out some more examples and see if you can figure out what is going on in each example: why do we need the extra premise for the inference to work?

These wildfires are out of control! So global warming is real.

The hidden assumption is something like "global warming is the best explanation for an increase in wildfires."

We should believe only what is reasonable, so we should reject theism.

The hidden assumption is something like "belief in theism is unreasonable."

No one believes the Earth is flat anymore, so it's a silly belief.

The hidden assumption here is something like "any belief that no one currently holds is a silly belief."

Here's a step-by-step process for identifying hidden assumptions:

A step-by-step process:

1. First, identify the inference or sub-inference with the hidden assumption

- Which one is "incomplete"?

2. Then, look at the premises of the inference and identify the "terms" or topics discussed in each premise

- Each premise is usually a claim which links two topics together.

3. Then, ask how we can link the terms that aren't yet linked.

- This requires a bit of imagination and instinct, but you can do it!

4. Finally, write the assumption that links the unlinked terms.

Now that you've identified a hidden assumption or more, perform the following two steps:

5. Check to be sure your argument now works

- Does the argument now have a link between each topic? Is there a path from the topics in the premises to the topics in the conclusion?

6. Perform the "negative test" on your assumption

- If you negate your hidden assumption, you should end up with an argument that makes no sense. If the argument with the negated premise makes sense, then you haven't identified a hidden assumption (i.e. the argument was fine without your assumption).

Let's take a look at how this works with some real arguments.

I think we should invade North Korea. Look, the Kim Jong dynasty is simply never going to give up on their goal of being a nuclear power.

Okay, this inference is really "fast" meaning that it skates over a few hidden assumptions and so doesn't seem to work all by itself. It's like it rushes straight to the finish line without actually running the course. Let's break it down step by step.

Step 1

There's only one apparent premise and one apparent conclusion, so identifying the inference in question is easy.

Step 2

The "terms" or topics of this inference are:

A. We should invade

B. North Korea

C. Kim Jong dynasty

D. Nuclear Arms

Step 3

How do we connect these topics? First, we need to connect "North Korea" with the "Kim Jong dynasty".

Then we'll need to connect "being a nuclear power" with "we should invade."

Step 4

Let's try these assumptions and see how the argument works out:

1. The **Kim Jong dynasty** is never going to give up on their goal of being a **nuclear power**.

2. The **Kim Jong dynasty** is going to lead **North Korea** for the foreseeable future.

3. Any country that aspires to be a **nuclear power** is one we should *invade*.

4. Therefore, we should *invade* **North Korea**.

This inference is more complete and connects the topics together more completely, but it rests on one very shaky premise. Can you identify which one?

Yes, that's right. You are very smart. Premise 3 is pretty wacky, right? What if Argentina decided it wanted to be a Nuclear Power? Should we invade them? I would hope not. They're a pretty harmless nation at the moment.

So, we have a choice. Either decide that the argument is pretty weak and reject it out of hand, or we can exercise the **Principal of Charity** to try to interpret the argument to be as plausible as possible. We should always interpret arguments—especially the ones we're skeptical of or disagree with—to be as rational and plausible as possible.

With that in mind, let’s change this argument up a bit so that it makes a bit more sense. We might not agree with the argument in the end, but at least we will have understood the argument in the best possible light. We will have seen what the most plausible argument for that conclusion on the basis of similar premises is.

If we ignore premise 3, the weak premise, and try to replace it with a few premises which make more specific and believable claims, we’ll be in a better spot. We’ll need to tie together some new topics. Premise 3 was supposed to make a link between “aspires to be a nuclear power” and “we should invade.” There’s a bit of conceptual “distance” between these ideas, though, so we shouldn’t just posit a principle like premise 3 above which directly links them. That would be too easy a principle to reject. Instead, we’ll travel the distance between these ideas in a few steps rather than a giant leap.

How about we connect “aspires to be a nuclear power” to “they’re dangerous”?

Then we can get from “they’re dangerous” to “we should invade.” That sounds more plausible, right?

1. The **Kim Jong dynasty** is never going to give up on their goal of being a **nuclear power**.

2. The **Kim Jong dynasty** is going to lead **North Korea** for the foreseeable future.

3. **North Korea** under **Kim Jong rule** would be an immediate existential **danger** to its neighbors and the rest of the world if they ever became a **nuclear power**.

4. If we *invade* **North Korea**, then we prevent that **danger**.

5. Therefore, we should *invade* **North Korea**.

Now the argument seems to hang together a bit more clearly. We have a clear path from the Kim Jong dynasty through to a nuclear North Korea, to the danger that poses and therefore a motivation for invading, all the way to the claim that we should invade. It’s still shaky reasoning, but it’s approaching the strongest version of the original argument.

There’s still technically something missing. Between 4 and 5 we’ve missed a premise.

In order to get from an “is” claim to an “ought” claim, often you’ll need a general *normative* principle. That is, we need a general rule which allows us to move from a simple statement of supposed fact (premise 4) to a prescription for what we *should* do (the conclusion, #5). This one will do:

4a. If we can prevent immediate existential danger to whole countries then we must/should act so as to prevent that danger.

That actually seems pretty plausible, right? So in this case the missing premise wasn’t so shaky (you might disagree, though).

We can then perform the negative test on our hidden assumptions and figure out if the argument falls apart without them. If we deny 4a, then we can’t get from 4 to 5. Does that make sense? If so,

Mapping Hidden Assumptions

Mapping Hidden Assumptions is relatively simple.

A hidden assumption will always offer **conjoint support** for its conclusion/sub-conclusion.

Think about it: if hidden assumptions are things that must be true for an inference to work, and conjoint premises are premises that must all be true for the inference to work, then it makes sense that any hidden premise will offer conjoint support.

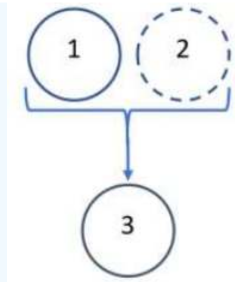
The only difference will be that we’ll use **dotted circles** instead of regular:

, or


, etc.

A few complete examples:

Example [Math Processing Error]



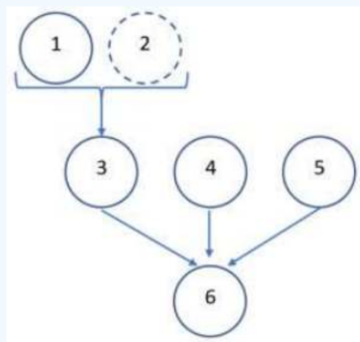
1. We have a right to bodily autonomy

2. Hidden Assumption: (Abortion restrictions infringe on a right to bodily autonomy)

3. Therefore, we have a right to freedom from abortion restrictions

The inference from 1 to 3 makes some sense because we're all familiar with the abortion debate by now. What we do to make it make sense of it for ourselves, though, is implicitly add in premise two in understanding the inference from 1 to 3.

Example [Math Processing Error]



(1) We'll never stop climate change, ***(2) Hidden Assumption: Climate change will intensify fires and storms,*** so (3) we're going to have much more fires and storms. (4) Our current system can't handle even basic disasters. Furthermore, (5) state disaster relief funds are insufficient without help from FEMA's overburdened funds. So, (6) we need to reform and enhance funding for FEMA immediately.

The inference from 1 to 3 makes little sense if it's not true that climate change is connected to fires and storms, so we need premise 2 to make that connection.

3, 4, and 5 are independent because each by itself makes sense as a premise for 6.

3.3: Objections

This is the final skill in argument mapping that we'll discuss. It's very important to be aware of because without it we can only ever map parts of a discussion which *agree* with one another. Meaning that if someone has an objection to an argument, we won't be able to pinpoint where the disagreement is on our argument maps. Our argument maps are more *expressive* and *powerful* if we can diagram disagreements as well as agreements, so let's discuss how to map disagreements!

We need to have a tool to map objections to an argument that we've already mapped so we can see where those objections are "putting pressure" or what exactly they're trying to critique.

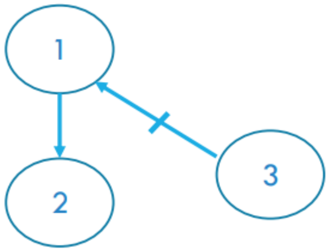
We use either hashed arrows or dotted arrows to signify relations of objection. They mean "objects to" or "rejects".

These are the only arrows in our system which point ***up.***

There are two kinds of objections:

1. Objections to a proposition

- Analyst A: (1) Building this highway would threaten the existence of the species of fairy shrimp that inhabits the proposed route, so (2) we must reroute the highway.
  Analyst B: I understand your concern for the fairy shrimp, but (3) the proposed route is not the only habitat for the fairy shrimp in this area, so it wouldn't threaten their existence.

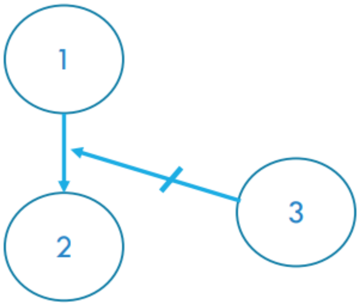


Note

Note that the arrow is pointing to the **Premise.**

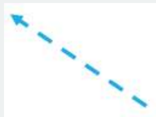
2. Objections to an inference

- Obama said that (1) people who were brought here illegally as children shouldn't be punished for the choices of their parents, because (2) their parents made the decision and not them. But that doesn't follow because (3) we're not punishing them, instead we're just enforcing our laws.

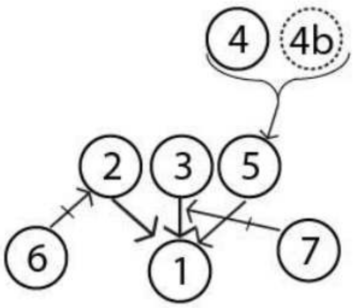


> **Note**
>
> Note that the arrow is pointing to the other **Arrow.**

Could have also used a dotted arrow.



Diagramming Tip



Here's a weird case, so that you can see what sorts of situations you might get into when trying to map objections:

Notice how objection 7 has to cross over an inference to point to the inference from 3 to 1. That's okay. But it helps to keep your argument maps a bit larger and as clear as possible. You want to be able to clearly decipher what's going on in the argument map as you go along.

3.4: More Complex Arguments

Here is a description of complex argument diagramming from Matthew Knachel.

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

The arguments we've looked at thus far have been quite short—only two or three premises. But of course some arguments are longer than that. Some are much longer. It may prove instructive, at this point, to tackle one of these longer bits of reasoning. It comes from the (fictional) master of analytical deductive reasoning, Sherlock Holmes. The following passage is from the first Holmes story—A Study in Scarlet, one of the few novels Arthur Conan Doyle wrote about his most famous character—and it's a bit of early dialogue that takes place shortly after Holmes and his longtime associate Dr. Watson meet for the first time. At that first meeting, Holmes did his typical Holmes-y thing, where he takes a quick glance at a person and then immediately makes some startling inference about them, stating some fact about them that it seems impossible he could have known. Here they are—Holmes and Watson—talking about it a day or two later. Holmes is the first to speak:

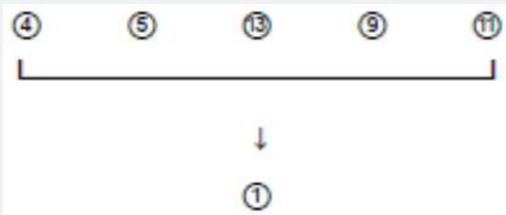
> "Observation with me is second nature. You appeared to be surprised when I told you, on our first meeting, that you had come from Afghanistan."
>
> "You were told, no doubt."
>
> "Nothing of the sort. I knew you came from Afghanistan. From long habit the train of thoughts ran so swiftly through my mind, that I arrived at the conclusion without being conscious of intermediate steps. There were such steps, however. The train of reasoning ran, 'Here is a gentleman of a medical type, but with the air of a military man. Clearly an army doctor, then. He has just come from the tropics, for his face is dark, and that is not the natural tint of his skin, for his wrists are fair. He has undergone hardship and sickness, as his haggard face says clearly. His left arm has been injured. He holds it in a stiff and unnatural manner. Where in the tropics could an English army doctor have seen much hardship and got his arm wounded? Clearly in Afghanistan.' The whole train of thought did not occupy a second. I then remarked that you came from Afghanistan, and you were astonished." (Also excerpted in Copi and Cohen, 2009, Introduction to Logic 13e, pp. 58 - 59.)

This is an extended inference, with lots of propositions leading to the conclusion that Watson had been in Afghanistan. Before we draw the diagram, let's number the propositions involved in the argument:

1. Watson was in Afghanistan.
2. Watson is a medical man.
3. Watson is a military man.
4. Watson is an army doctor.
5. Watson has just come from the tropics.
6. Watson's face is dark.
7. Watson's skin is not naturally dark.
8. Watson's wrists are fair.
9. Watson has undergone hardship and sickness.
10. Watson's face is haggard.
11. Watson's arm has been injured.
12. Watson holds his arm stiffly and unnaturally.
13. Only in Afghanistan could an English army doctor have been in the tropics, seen much hardship and got his arm wounded.

Lots of propositions, but they're mostly straightforward, right from the text. We just had to do a bit of paraphrasing on the last one—Holmes asks a rhetorical question and answers it, the upshot of which is the general proposition in 13. We know that proposition 1 is our conclusion, so that goes at the bottom of the diagram. The best thing to do is to start there and work our way up. Our next question is: Which premise or premises support that conclusion most directly? What goes on the next level up on our diagram?

It seems fairly clear that proposition 13 belongs on that level. The question is whether it is alone there, with an arrow from 13 to 1, or whether it needs some help. The answer is that it needs help. This is the general/particular pattern we identified above. The conclusion is about a particular individual—Watson. Proposition 13 is entirely general (presumably Holmes knows this because he reads the paper and knows the disposition of Her Majesty's troops throughout the Empire); it does not mention Watson. So proposition 13 needs help from other propositions that give us the relevant particulars about the individual, Watson. A number of conditions are laid out that a person must meet in order for us to conclude that they've been in Afghanistan: army doctor, being in the tropics, undergoing hardship, getting wounded. That Watson satisfies these conditions is asserted by, respectively, propositions 4, 5, 9, and 11. Those are the propositions that must work jointly with the general proposition 13 to give us our particular conclusion about Watson:



Next, we must figure out how what happens at the next level up. How are propositions 4, 5, 13, 9, and 11 justified? As we noted, the justification for 13 happens off-screen, as it were. Holmes is able to make that generalization because he follows the news and knows, presumably, that the only place in the British Empire where army troops are actively fighting in tropics is Afghanistan. The justification for the other propositions, however, is right there in the text.

Let's take them one at a time. First, proposition 4: Watson is an army doctor. How does Holmes support this claim? With propositions 2 and 3, which tell us that Watson is a medical and a military man, respectively. This is another pattern we've identified: these two proposition jointly support 4, because they each provide half of what we need to get there. There are two parts to the claim in 4: army and doctor. 2 gives us the doctor part; 3 gives us the army part. 2 and 3 jointly support 4.

Skipping 5 (it's a bit more involved), let's turn to 9 and 11, which are easily dispatched. What's the reason for believing 9, that Watson has suffered hardship? Go back to the passage. It's his haggard face that testifies to his suffering. Proposition 10 supports 9. Now 11: what evidence do we have that Watson's arm has been injured? Proposition 12: he holds it stiffly and unnaturally. 12 supports 11.

Finally, proposition 5: Watson was in the tropics. There are three propositions involved in supporting this one: 6, 7, and 8. Proposition 6 tells us Watson's face is dark; 7 tells us that his skin isn't naturally dark; 8 tells us his wrists are fair (light-colored skin). It's tempting to think that 6 on its own—dark skin—supports the claim that he was in the tropics. But it does not. One can have dark skin and not visited the tropics, provided one's skin is naturally dark. What tells us Watson has been in the tropics is that he has a tan—that his skin is dark and that's not its natural tone. 6 and 7 jointly support 5. And how do we know Watson's skin isn't naturally dark? By checking his wrists, which are fair: proposition 8 supports 7.

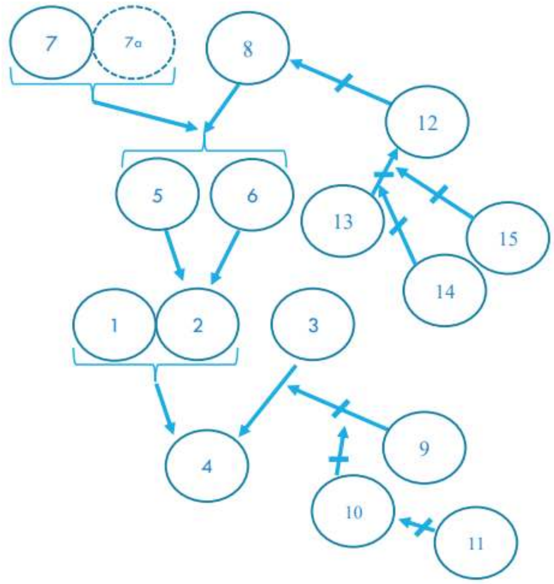
So this is our final diagram:

And there we go. An apparently unwieldy passage—thirteen propositions!—turns out not to be so bad. The lesson is that we must go step by step: start by identifying the conclusion, then ask which proposition(s) most directly support it; from there, work back until all the propositions have been diagrammed. Every long argument is just composed out of smaller, easily analyzed inferences.

3.5: Argument Mapping Conclusion

Theoretically, you could have a very complex argument map which traces lines of disagreement through multiple stages. We have a sophisticated enough argument mapping system now that we would be able to map even absurdly complex disputes.

As a little practice, try to identify each component here and then try to figure out which numbers would belong to which of two people in a dispute that has this map.



3.6: Beginning to Evaluate Arguments

Before I said that we use argument mapping to understand an argument before we do any evaluation of those arguments as good or bad arguments, valid or invalid, cogent or uncogent, sound or unsound, and so on.

But the process of identifying hidden assumptions is in itself a sort of evaluative process: we must identify the *need* in the argument for another claim to be true—we have to declare that an argument is incomplete and therefore faulty before we can talk about the argument *needing* an extra premise.

So if we've already done a bit of evaluating of arguments in identifying hidden premises, perhaps those tools that we gained when we were identifying hidden assumptions will help us in trying to evaluate an argument as a good or bad argument. Let's explore how helpful those tools are.

Recall that the central idea in the process of identifying hidden assumptions is the idea that a good, complete argument has a series of topics or terms that are linked together in the right way. This is a really informal and incomplete version of what you might learn if you learn Categorical or Aristotelian Logic. That's the study of how categories or terms must be linked for an argument to be deductively valid.

Let's look at how this informal idea can help us determine whether we've got a good or bad argument. Here's an example to consider—it's something you might see on social media and in the wording that one might actually find there:

If you don't want an abortion, don't have one.

[Implied conclusion: Abortion should remain legal and available].

Hopefully it's relatively clear that this argument needs a tune up—not because it is wrong-headed (it might not be) or because its conclusion is wrong (it might be true), but because it's unclear what the actual argument is: what are the premises and how do they support the conclusion?

First, we should try to interpret the argument in a way that makes the process of identifying topics easier. This process is already a bit evaluative in that we are interpreting the argument charitably: we're trying to understand the most rational version of the argument without changing its essential content. Maybe something like:

Abortion rights aren't abortion mandates,

so there's no reason to oppose abortion rights.

Making abortion legal doesn't mean requiring (mandating) anyone to have an abortion. This is already starting to look more complete, and it's already beginning to look like a serious philosophical argument. We've taken it out of the format of a slogan and interpreted it as an argument.

The next step is to identify the topics being discussed:

a. Abortion rights laws

b. Abortion mandates

c. Reason to oppose things

Once we've identified these terms, we can think about the argument in terms of a series of propositions which connect them together.

1. Abortion Rights laws aren't abortion mandates

[Links (a) to (b)]

2. So, There's no reason to oppose abortion rights

[Links (c) to (a)]

We've got some of the tools now to recognize that there's a missing premise here. Can you find it?

Ummm....There's no link between Reason to oppose things and Abortion mandates, so the inference has a gap in it.

Right-o! Good work. We've linked a to b, and a to c, but not b to c. This is why the argument feels a bit gappy and incomplete. What, then, might our hidden premise be? We need to link "Abortion mandates" with "reason to oppose things." Any ideas?

Yeah. Why not something like "There's no reason to oppose abortion mandates."

Well...not quite. That's a pretty clearly false, statement right? Lots of people *want* to have children and lots of other people are morally opposed to abortions, so it makes little sense to say there's no reason to oppose mandating everyone has an abortion, right?

Oh...yeah. What about "If something isn't a mandate, then there's no reason to oppose it."

Now we're talking! That's a really general claim: it applies to everything (or at least every public policy or law) that isn't a mandate. So maybe we'd want something more restrictive. Nevertheless, this is a good start. So here's our complete argument:

1. Abortion rights laws aren't abortion mandates

[Links (a) to (b)]

1a. If something isn't a mandate, then there's no reason to oppose it

[Links (b) to (c)]

2. So, There's no reason to oppose abortion rights laws

[Links (c) to (a)]

This process is tricky and interpretive. When we get to Aristotelian Logic, we'll cover some tools that make this a more strict and formalized process. For now, though, we're sticking with intuitively understanding arguments and how they hang together. This is an incredibly valuable skill in all aspects of life: when is an inference making an implicit assumption? People make loads of implicit assumptions all of the time.

Okay, so the goal in this section was to start to evaluate arguments using some of the tools we've picked up so far. Here's how this might go:

The first claim in this argument is that **Abortion rights aren't abortion mandates**. This is clearly true. No law or policy securing abortion rights (within the realm of reasonable laws) would *require* that people have abortions. So there's no problem with the first claim.

The second (hidden) claim in the argument is that **If something isn't a mandate, then there's no reason to oppose it.** This, as I said before, is *very* broad. It seems easy to come up with a counterexample. In this case, a counterexample would be something that isn't a mandate, but which we would have reason to oppose. Can you think of a possible law that doesn't put a requirement on citizens, but nevertheless is a law we'd want to oppose?

I can think of thousands. A law allowing murder. A law allowing ritual human sacrifice. A piece of legislation requesting that the President strongly consider nuking the moon....the list goes on.

So this seems like a bad principle: one shouldn't be opposed to something that doesn't bring a mandate with it. In other words: if it doesn't interfere with you, then leave it alone. That's not how a civil society, works though. We take an interest in each other's lives for the sake of having a healthy society and protecting one another.

As I said a few paragraphs back: this hidden premise we inserted completes the argument fairly directly, but might be *too general*. More general than we need, in fact. We might consider something more specific like "If a law doesn't require people to do something immoral, then there's no reason to oppose it." But that is refuted by the same counterexamples I listed above. We could try over and over to find more specific premises to complete the argument, but honestly at some point we have to admit that the argument as we've reconstructed it probably rests on a false general premise.

So in reconstructing this argument and identifying what hidden premise it rests on, we have set ourselves up to evaluate whether or not this is a good argument. You might disagree, but this seems like a bad argument for abortion rights. Best to go looking elsewhere for good arguments for abortion rights (they're out there! Philosophers like Judith Jarvis Thomson, Mary Anne Warren, and Rosalind Hursthouse have interesting arguments indeed).

3.7: Chapter 3 - Key Terms

- **Conjoint Support**
- **Independent Support**
- **Hidden Assumption/ Suppressed Premise**
- **Sub-Inference**
- **Sub-Conclusion**
- **Sub-Premise**
- **Main Inference**
- **Main Premise**
- **Main Conclusion**
- **Objection to a Premise**
- **Objection to an Inference**

3.E: Chapter Three (Exercises)

How can I tell one from the other?

- Objections to inferences usually use phrases like "that doesn't follow," or "even if we accept x, we need not accept y."
  - The objection is that the inference itself is incomplete or weak or invalid or simply doesn't make any sense.
  - An objection to a hidden premise is actually an objection to an inference: you're claiming that the inference rests on a weak hidden premise and so is an incomplete inference.
- Objections to propositions/premises will always be arguing that some proposition is false rather than that a conclusion doesn't follow.
  - The objection is that some particular claim is false or at least is likely or plausibly false.

Exercise $3.E.1$: Conjoint vs Independent Support

For each, determine whether the support offered by the premises is conjoint or independent. Deploy the negative test when you're unsure.

A. Eating healthy food is important and Figs are super healthy, so we should eat more figs.

B. I have to have a steady income to support my family, I already have a stable job, and grad school would require me to quit my job, so I shouldn't go to grad school.

C. All of the nurses have gone on the strike, the custodial staff is threatening the same, and the doctors are demanding better legal support. This hospital is in trouble right now.

D. He is ten years younger than you and no one should date anyone ten years younger, so you can't date him!

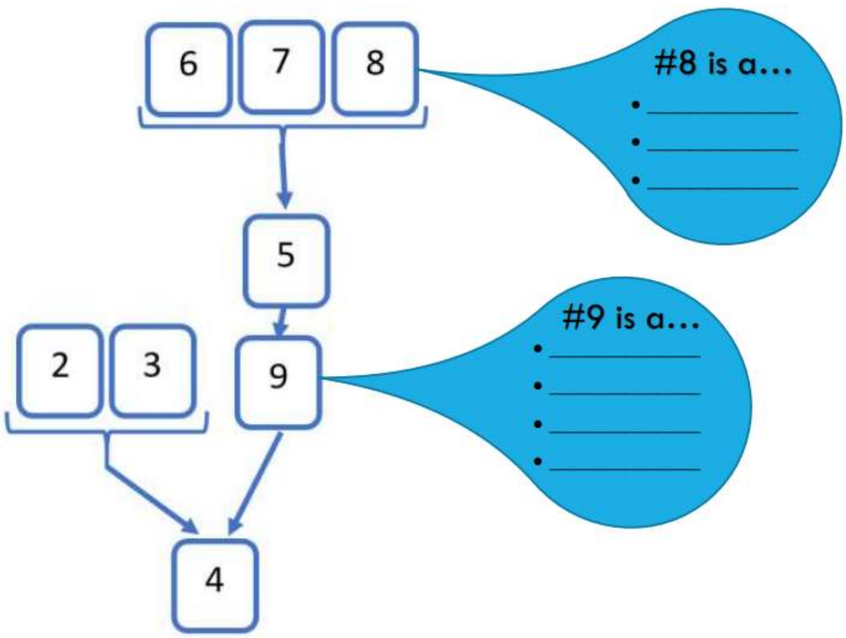
E. An ergonomic desk can prevent permanent injury, is more comfortable to use, and is cost-effective, so can I please buy one for my office?

F. A robust economic recovery will require higher taxes on the wealthy, and we need to have a strong recovery to prevent melt downs in the near future, so we must raise taxes on the wealthy.

G. We've always been honest with each other and the honest thing to do right now is to tell you that that outfit is terrible, so I need to tell you the truth about that outfit.

Exercise $3.E.2$: Terminology

Fill in the blank labels using one of each of the following key terms: (A) Sub-Conclusion, (B) Sub-Premise, (C) Main Premise, (D) A premise on level 4 of the argument map, (E) A premise on level 2 of the argument map, (F) Conjoint Premise, (G) Independent Premise.



? Exercise $3.E.3$: Simple Argument Maps

For each, create an argument map. Be sure to distinguish between conjoint and independent support.

A. (1) Eating healthy food is the most effective weight-loss strategy, since (2) the amount of calories one takes in while eating even small snacks takes a long time to burn off by exercising, and (3) almost no one can afford to spend hours and hours exercising throughout the day.

B. (1) We've been out here in the sun all day, and (2) being in the sun for too long is unhealthy, so (3) let's go inside.

C. (1) He's so popular. (2) Everyone wanted to be invited to his birthday party and (3) he had five people invite him to the dance.

D. (1) We should ban all guns. (2) Guns are especially effective killing machines for mass killings. Also, (3) children often have fatal accidents with guns. Furthermore, (4) guns don't have a non-violent use.

E. (1) Having intercourse before 18 is wrong because, (2) you are not emotionally mature enough to deal with the awkwardness and intimacy of the situation, and (3) you are not mature enough to deal with the potential consequences of the situation (pregnancy or STIs).

F. (1) Treating others with respect is important, so (2) we should all respect each other, and (3) we should try to teach our children to respect others.

G. (1) Oreos aren't healthy, since (2) Nabisco products generally aren't healthy. Think about it: (3) Pringles aren't healthy, (4) Pop Tarts aren't healthy, and (5) neither are Chips Ahoy.

H. (1) People are starting not to like you. (2) Tina said she wasn't your friend anymore, (3) Beto said he doesn't like you, and (4) I certainly don't want to be around you.

? Exercise $3.E.4$: More Complex Argument Maps

For each, create an argument map. Be sure to distinguish between conjoint and independent support.

A. (1) I know that Sally went to the park with Billy because (2) Sally said she'd go with him if he asked, and (3) Billy likes Sally (so he wouldn't ask her as a prank) and (4) Billy asked Sally to go to the park.

B. (1) Eating any meat is wrong because, (2) most meat is produced in factory farms, (3) animals in factory farms suffer greatly, and (4) even 'free range' and 'organic' meat causes animal suffering.

C. (1) We already have almost all of the technology needed to clone dinosaurs and (2) human beings tend to do whatever they find they *can* do. (3) So, killer dinosaurs will roam the Earth one day. And since this is true, we can expect two things: (4) an armed response leading to the loss of innocent life, and (5) movie producers trying to buy the rights to the story.

D. (1) Burning fossil fuels like petrol, coal, and natural gas contributes to global warming. (2) According to experts, the combustion reaction releases free molecules of CO_2 into the atmosphere, and (3) Scientists wouldn't lie about this. Think about it: (4) there's no profit incentive for scientists to lie about this, but (5) there *is* a profit motive for other people to deny that it is true.

E. The Republicans have argued repeatedly that (1) the Affordable Care Act is in a death spiral. Because, they say, (2) premiums are getting higher, and (3) as premiums get higher, the people will stop purchasing policies and (4) if the people stop purchasing policies, then the insurance companies will pull out of the exchanges, and (5) if that happens, then the whole system collapses.

F. (1) We need to protect the environment, since (2) biodiversity is necessary to protect future food sources and (3) biodiversity is sustainable only in a relatively healthy global environment. Furthermore, (4) We take great pleasure in the natural wonders that the Earth has to offer (5) [suppressed premise]

? Exercise $3.E.5$: Even More Complex Argument Maps

For each, create an argument map. Be sure to distinguish between conjoint and independent support.

A. (1) We need to buy a new trampoline, since (2) our son almost hurt himself really badly when this one broke last week, and (3) I don't want to risk it again. (4) Even if you're able to fix it, there's no guarantee that it will be as safe as a new one. Think about it: (5) older trampolines like ours don't have a net around them, and (6) the net makes it less likely that a kid will bounce off onto the ground and hurt themselves. Finally, (7) older trampolines like ours don't have good spring covers, and (8) without spring covers, the risk of pinching oneself or falling through the springs and breaking a limb are very high.

B. (1) Eating kale is sometimes unsatisfying, but the fact is that (2) Kale has countless health benefits. (3) It is rich in folate and (4) folate helps guard against bad epigenetic changes. (5) It has more minerals and vitamins than most meat sources and (6) vitamins and minerals got from whole food sources are better than those got from multivitamins and other supplements since (7) whole foods contain more bioavailable forms of vitamins and minerals.

C. (1) We've already been in Afghanistan for over a decade and (2) no other American war has lasted this long, so (3) Afghanistan is the longest running American war. (4) We've shown little sign of progress in the past few years, and (5) we've sunk countless dollars into Afghan infrastructure and security projects with little to show for it. Given all of this, (6) we should pull out of Afghanistan and (7) we should divest interest in the Afghan society. (8) Since we've already tried so hard to fix it, (9) we should let them try to solve their own problems!

D. (1) Epigenetics is the most important frontier in genetic research. (2) Countless traits and processes depend not on genetic changes, but on epigenetic changes, (3) epigenetic changes are easier to induce through therapies, chemicals, and other interventions in a clinical setting, and (4) we already know the basic rules of genetics, but are far behind in our understanding of epigenetics. Given all that, it follows that (5) we should shift the balance of funding in favor of epigenetic research and (6) we should fund more PhD's in epigenetics as well.

E. (1) We should put more direct emphasis in school and college on thinking clearly and critically. (2) The most important skill in life is thinking well. I think this because (3) other important skills like decision making and communication rely centrally on thinking well, and (4) a good citizen, employee, and overall person is one who can think clearly and rationally. (5) Citizens must weigh complex values in voting on candidates and referenda, (6) employees must make decisions in the workplace based on complex policies and competing needs, and (7) people in general need to have habits of self-critical and careful thinking in order to live good lives.

? Exercise $3.E.6$: Hidden Assumptions

For each inference, identify the most direct hidden assumption.

A. Moby Dick is a whale. So Moby Dick is a mammal.

B. Giving students a fail grade will damage their self-confidence. Therefore, we should not fail students.

C. It should not be illegal for adults to smoke pot. After all, it does not harm anyone.

D. There is nothing wrong with texting during lectures. Other students do it all the time.

E. Traces of ammonia have been found in Mars' atmosphere. So there must be life on Mars.

F. I don't like people who spit on the sidewalk, so littering should be illegal.

G. No one even cares what you think, so what you think isn't important.

H. Americans believe in freedom, so any law that restricts our freedom should be abolished.

I. Trees are beautiful, so we should plant more of them.

J. Carbon emissions contribute to global warming, so we should tax them.

Exercise $3.E.7$: Mapping Hidden Assumptions

For each inference, identify the hidden assumption and then create a map of the inference including the hidden assumption.

A. The truth is, (1) we can't vote for the Republican candidate. (2) She doesn't believe in global warming.

B. (1) Nobody has ever been there and come back, and (2) I have children, so (3) I'm not going.

C. (1) Freedom isn't free. (2) So, someone has to pay the price for freedom. (3) The way people pay the price of freedom is by serving in the armed forces. (4) So we should institute a draft. [at least two hidden assumptions]

D. (1) Nobody has ever seen a dinosaur, so (3) dinosaurs don't exist.

E. (1) We should reduce the penalty for drunken driving, as (2) a milder penalty would mean more convictions. (3) The only way to reduce the penalty is to elect more liberal judges and prosecutors, so (4) we should elect liberal judges and prosecutors.

F. (1) Never again should we bow to tyrants, because (2) tyranny has been the mark of rule throughout human history, (3) as has cruelty and abject want. It follows that (4) we must rebel against the Imperial rule of England.

G. (1) Only real marriages should be recognized by the state, so (2) polygamist marriages shouldn't be recognized by the state. (3) Any marriage not recognized by the state should be illegal. So (4) polygamist marriages should be illegal. (5) Another reason they should be illegal is that, polygamist marriages often result in abusive situations. [the hidden assumption is between 1 and 2]

H. (1) No one believes in Odin anymore, so (2) why should anyone believe in God? [this is a rhetorical question, which is a claim that is disguised as a question. The claim appears to be "no one should believe in God"]. (3) If no one should rationally believe in something, then we should actively fight against belief in it. It follows that (4) we should actively fight against belied in the existence of God. (5) A world without believers would be a better world to live in. [the hidden assumption is between 1 and 2]

I. (1) We can't let terrorists live here with us in Pakistan, so (2) we should expel all Christians from our country. (3) Christians also don't contribute to the economy and (4) could potentially be spies for the Americans. [Where's the most blatant hidden assumption? There are more than one, but one in particular is relatively clearly a missing assumption of the argument]

Exercise $3.E.8$: Identifying Types of Objections

Identify which type of objection is illustrated: an objection to a premise or an objection to an inference (including pointing out that there's a hidden premise and/or rejecting a hidden premise)?

A. I agree with your conclusion, but it doesn't follow from your assumptions.

B. Interesting argument, but what I don't understand is your claim that every case of tyranny is a case of injustice. That doesn't seem quite right.

C. You claim that there isn't a threat to the Amazon. On the contrary, there are countless threats, one of which is people claiming that there isn't a threat to the Amazon!

D. So if I accept all of your assumptions, it doesn't seem to me that I must accept your conclusion.

E. If I have it right, it seems to me that your inference rests on a hidden assumption that we ought to do whatever is in our national interest. That's not clearly true. Think about cases of humanitarian aid that only very indirectly if at all are in our national interest.

F. I think I understand the general thrust of this argument, but one claim makes me uncomfortable. Your inference rests on the claim, as you stated it, that Great Britain is to blame for more historical atrocities than any other European nation. That's not clearly right.

G. I don't think this is a good argument. We won't clearly advance well beyond where we are today in terms of computing power because of the physical limits of the hardware we have available.

? Exercise $3.E.9$: Mapping Ojections

Identify which kind of objection is illustrated and then map the objection along with the original argument.

A. Person A: (1) Edward Snowden released petabytes of classified data. (2) He should be convicted of treason.

Person B: Wait a minute! (3) We shouldn't just convict anyone who releases that much data of treason!

Person A: (4) If we don't, then we'll be opening the door to more dangerous leaks.

Person B: (5) Actually, come to think of it, I don't think he did release petabytes. I think it was only Terabytes.

B. The Republicans have argued repeatedly that (1) the Affordable Care Act is in a death spiral. Because, they say, (2) premiums are getting higher, and (3) as premiums get higher, the people will stop purchasing policies and (4) if the people stop purchasing policies, then the insurance companies will pull out of the exchanges, and (5) if that happens, then the whole system collapses. But their conclusion doesn't follow, since (6) people need health insurance and won't stop purchasing it if prices continue to rise incrementally.

C. Person A: (1) College isn't designed around the goal of producing good plumbers and electricians and welders. (2) Furthermore, college is expensive and (3) college is time-consuming. So (4) we shouldn't expect everyone to go to college.

Person B: I understand your inference, but (5) college does make one a better plumber, electrician, and welder because it gives you a host of intellectual resources to bring to bear on solving the many unforeseen problems that arise on jobs like that.

Person C: I actually take issue with the inference here from your first claim to your conclusion, since (6) college isn't about job training, but is instead about creating a well-informed citizenry that can make rational and informed decisions at the voting booth.

D. Obama argued that (1) we should pass the ACA, claiming that (2) there is an epidemic of chronically-ill citizens without health insurance due to their pre-existing conditions and that (3) many citizens simply can't afford health insurance.

But (4) the ACA won't provide health insurance to a large group of relatively poor Americans.

E. Her argument was as follows: "(1) No one wants to be put in the position where they are faced with a deadly intruder without the proper means to protect theirself and their family. (2) Gun laws make it probable that someone will end up in that situation. (3) Therefore, we can't enact gun control legislation."

But that argument isn't convincing. (4) Even if we accept the premises, we need not accept the conclusion. After all, there are reasons to pass gun control that must be addressed.

? Exercise $3.E.10$: Hidden Assumptions and Objections

Identify the hidden assumptions in the first argument and then map both the argument and the objections. Remember that objections to hidden assumptions are objections to *inferences* and so they should be mapped as such.

A. Frank: (1) We'll never make it to the party on time, so (2) let's just turn around and head home. (3) Samir and Imani live miles away and (4) we can't go very fast in this traffic.

Margaret: That's ridiculous, we'll absolutely make it on time. First, (5) we have 30 minutes to get there and also (6) we could be 15 minutes late and still be "on time" since it's a party.

B. Tamil: (1) We need to protect the environment, since (2) biodiversity is necessary to protect future food sources and (3) biodiversity is sustainable only in a relatively healthy global environment. Furthermore, (4) we take great pleasure in the natural wonders that the Earth has to offer (5) [suppressed premise]

Jamal: (6) I agree with your conclusion, but even if we accept that biodiversity is necessary and that protecting the environment is necessary for protecting biodiversity, we need not accept your conclusion.

C. (1) Counting Crows wrote and performed Mrs. Robinson, so (2) They're the best band ever.

Ummm... (3) they wrote and performed "Mr. Jones", not Mrs. Robinson. And either way (4) neither song would make them the best band ever.

D. He said "(1) I need some space, so (2) we need to break up." But (3) he doesn't need space. And either way, (4) needing space isn't a good enough reason to break up with someone.

E. She said "(1) Potato chips are high in saturated fat and salt, and so (2) they should be consumed very sparingly." But that's a bad inference since (3) dietary research is overturning the idea that saturated fat is bad for humans and (4) humans need salt to maintain proper blood volume and electrolyte concentrations.

F. Pablo: (1) We shouldn't eat even fake animal meat since (2) we wouldn't think it's okay to eat fake human meat. Afterall, (3) eating fake human meat would be tacitly affirming that cannibalism is morally acceptable. Marisela: I disagree, (4) there's a faulty hidden premise there: that eating fake animal meat is analogous to eating fake human meat. Furthermore, (5) the other inference for the claim that eating fake human Meat is wrong has a hidden assumption as well and I'm not so sure it's correct.

CHAPTER OVERVIEW

4: Intellectual Virtues and Vices

Michael Fitzpatrick contributed quite a bit to this chapter, so Chapter 4 should be seen as a collaboration between Lavin and Fitzpatrick.

Note for Instructors

Starting in edition 4, Chapter 4 on fallacies has been renamed and somewhat rewritten into a chapter on "Intellectual Virtues and Vices," incorporating material from the conclusion and the previous version of the chapter. Thinking about the traditional fallacies under a model of virtue epistemology seemed more in line with the values of this textbook, but they can still be taught as traditional fallacies if that is your preference.

The most significant change is that the "Fallacies of Induction," which was 4.3, have all been moved to the new 8.5, forming part of the chapter on inductive reasoning. This allows those fallacies to be taught in the context in which they are mistakes in reasoning, as teaching fallacies in the context of their positive counterparts seems more pedagogically useful.

Also, the fallacy of equivocation has been moved to 2.2 "Fallacy of Equivocation", since it is most naturally taught alongside the discussion of the role of language in critical thinking.

All the other fallacies (Relevance and Presumption) remain here, described as intellectual vices.

4.1: What are Virtues and Vices?

It can be tempting to suppose that thinking well is simply a matter of following the right rules. As long as we evaluate our evidence in the right way, make sure our premises support our conclusions, and don't make mistakes, then we'll be good thinkers.

But this actually isn't the case. As much as we might wish, good thinking is not ultimately about following rules. It can seem that way, especially when we get to our later chapters on deductive logic, but in reality only a small number of situations are appropriate for deductive reasoning. We learn rules of reasoning primarily to train our brains how to think clearly and carefully, not because the rules will be always applicable. In most situations thinking well is a matter of good judgment, where we have to decide what makes the most sense to believe given *these* particular facts and values in *this* particular situation. How we reason in one context might not make sense in the next context when new information or methods of investigation arise.

We learn how to map arguments and understand validity to discover basic patterns that generally lead towards truthful beliefs. Thinking well means learning when particular patterns apply and when they don't. So how do we become people of good judgment?

Good judgment requires combining our reasoning abilities and the techniques we learn in this class with a practice of building up in ourselves intellectual **virtues** while avoiding intellectual **vices**. Virtues are character traits or dispositions about a person that help them be a good overall person. Artistic virtues make one a good artist; social virtues make us likeable to others, and ethical virtues help us to promote flourishing in our own lives and the lives of others. The intellectual virtues are like these—they help us be better thinkers and to think well with others. It's not just how we think that matters; it also matters the kind of person we are.

Examples can help, so let's take a quick glance at some artistic virtues to help us understand what we're talking about. One artistic virtue is probably *creativity*. Artists must be creative people, who can take familiar representational materials and imagine new, purposeful ways to create those materials and present their creations as art. It's difficult to flourish as an artist if one lacks creativity. Good desires are also important virtues; an artist who does not *desire to create* will find it difficult to employ their creativity. It's not enough to have a creative mind; you also have to be motivated to use it. Finally, creating a piece of art, whether a painting or a collage or a sculpture or a theatre production, is hard work and takes an enormous amount of *patience* and *perseverance*. Creative people with a desire to create can still fall short of their artistic ambitions if they don't have the patience and perseverance to see their project all the way through.

I mentioned there are social virtues (sometimes called "social graces") and ethical virtues. Can you think of what some of these might be, using the artistic examples as a guide?

Of course, our textbook is about thinking well, and our focus will be on the intellectual virtues and their vices. Learning some intellectual virtues uses our character as thinkers to explain when we think well and when we don't. To keep our topic manageable, we're only going to focus on four central intellectual virtues, even though there are many, many more. Then, we'll discuss some ways people can lack virtue in their thinking, what we'll call intellectual vices. Vices are character traits or dispositions which inhibit our flourishing—so intellectual vices are those that make us think worse, rather than well.

A Word of Caution

The skills you'll pick up in this chapter—skills in identifying virtues and vices—can often be used as weapons. Especially online, where the goal is often to win and even humiliate rather than to connect and understand, charging someone with lacking virtue can be treated as a way of shutting someone out of a conversation. Don't use them this way.

The primary goal of learning how reasoning goes wrong is always to learn to think more clearly and to better ***yourself***. When these tools are used to make you seem more worthy of having your voice heard, they are being misused. So instead of being on the lookout for bad reasoning in others and being quick to shout "VICES!!" when someone missteps in their reasoning, instead be sure that you're having the discussion you're having because you want to understand the viewpoint of another, and take great care with how you treat those who haven't had the privilege of taking a class like logic and critical thinking with an amazing teacher like yours ;).

In short, focus on your own reasoning, but when you feel you must educate someone else, do so gently and in a spirit of mutual understanding.

4.2: Some Intellectual Virtues

Curiosity

As we mentioned in 1.1 "On Truth", truth matters because it satisfies our curiosity. But that means we need to be curious in the first place! Good thinkers have a disposition to pursue the truth for its own sake, solely out of a desire to know and not simply for the potential advantages it might afford us. Having a curious disposition means you are motived by a desire to learn about the world around you. You genuinely care about, say, the experience of black folk in urban housing, or whether a messenger mRNA molecule can be used in vaccines to produce an immune response. Of course, in both of these examples what we learn might turn out to be quite useful for improving the lives of minority communities or protecting human communities generally from harmful viral infections. But if we're curious, we will value learning about these things even if they don't lead to useful outcomes, simply because discovering the truth about our world enriches our lives all by itself, and learning more about how different communities struggle to create spaces to call home, or how various biochemical molecules work, enrich our lives whether or not we can develop policies or technologies as a result.

Being motivated to learn for its own sake is important for how we think well with others. If all I care about is what my learning will get me, I will be tempted or biased to prefer evidence that favors the outcomes I want, or to discount evidence for the outcomes I oppose. Being curious helps to guard us against caring about truth only for its instrumental value, rather than its intrinsic value. Curiosity places our focus on the question or puzzle we're interested in, and *not* on preserving a pre-determined position we already believe. The same goes for listening to other people. If I'm curious about how others see the world, it helps me to listen to their ideas and perspectives regardless of whether I can use what I learn to further my career or take advantage of them. Curiosity also helps us guard ourselves against manipulation; if we genuinely want to understand current events, we will be less interested in partisan or ideological news media and political commentary that "spins" events to serve a particular agenda.

Finally, curiosity is important as a virtue because it reminds us that making mistakes and being wrong is okay. The British-American philosopher Alfred North Whitehead wrote, "panic of error is the death of progress; and love of truth is its safeguard."[1] Making mistakes and judging falsely are not bad as long as we see them as part of the larger learning process. Mistakes give us an opportunity to figure out why we were mistaken, and make corrections accordingly. If we are curious, that is, if we are people motivated by a love of truth, then we will care more about the discovery process than the fact that we believed the wrong things along the way.

Intellectual Honesty

Intellectual honesty is the disposition to be truthful and sober in your assessment of your own knowledge. It's easy to claim that we know things and even to have confidence in what we know, but often we find that on reflection we shouldn't have as much confidence as we do. Confidence is cheap. What is of higher worth is the ability and disposition to recognize the things we don't know or shouldn't be confident in and the things that we do know and do have reason to be confident in. Much of what we think we know we think we know really because we read a headline while scrolling through Facebook or Twitter or someone told us once sort of off-handedly. These, when we think about it, aren't very good sources of knowledge. They aren't really grounds or justifications for our beliefs—or at any rate aren't very good justifications for our beliefs. Intellectual honesty is the disposition to take a beat, think about why it is that we feel confident in a belief and feel ready to assert it, and then proceed with a more honest assessment of what we know and why we think we know it.

Intellectual Humility

Intellectual humility goes hand in hand with intellectual honesty. What is means to be intellectually humble, though, is slightly different from being honest. Intellectual humility is a disposition to recognize that even when we have good grounds for knowing something, there might always be something that upsets that understanding or set of beliefs. To be intellectually humble is to remember that human beings have been very confident many times in the past and often for very good reason, but have turned out to be wrong due to some false assumption somewhere in their thinking. It's the disposition to say "even if I have really good reason to believe what I believe, I still might be wrong."

The Search for Vulcan

In the early 1800's, several astronomers including Alexis Bouvard noticed that the planet Uranus was not orbiting the sun in a manner consistent with current mathematical models about the laws of nature governing how planets move. This led Bouvard and another astronomer, Urbain Le Verrier, to postulate that there must be another planet in the vicinity whose gravitational pull was affecting the motion of Uranus. In 1846, using predictions sent to him by Le Verrier, Johan Gottfied Galle was able to spot a planet from the Berlin Observatory, which would become known as the planet Neptune.

Shortly after the discovery, Le Verrier turned his attention to Mercury, another planet that astronmers had trouble applying current physical predictions to. Le Verrier made a complete model of Mercury's motion with predictions to be tested when Mercury was next scheduled to orbit across the face of the Sun in 1848. Mercury failed to move in accordance with Le Verrier's predictions.

Rather than give up, Le Verrier spent the next decade creating some of the most rigorous and detailed calculations of the motion of Mercury to date, yet he could not come up with predictions that matched observation. Using inspiration from the successful prediction of Neptune as a planetary body affecting the motion of Uranus, Le Verrier predicted that there must be a planetary body affecting Mercury, and he postulated the existence of the planet Vulcan (same word that is used in some of the Star Trek stories!).

Over the rest of his life, Le Verrier worked with observatories to confirm the existence of Vulcan, and while many alleged sightings were reported, none came in that could be confirmed. He died in 1877, firmly believing that Vulcan was out there. It wouldn't be until 1915 that astrophysicists would finally be able to show that, in fact, there is no planet Vulcan—the behavior of Mercury can be explained by the curvature of spacetime, which was Albert Einstein's new way to account for the effects of gravity. Einstein's new theory correctly predicted the orbit of Mercury.

Notice that Le Verrier was motivated in his postulates of Neptune and Vulcan to satisfy his curiosity. The existence of a new planet at that time would have almost no technological significance. He simply wanted to know why there were small deviations in the Mercurial orbit where there should not have been. Using classical Newtonian mechanics, a nearby object's gravitational pull seemed like the most likely hypothesis. Since this hypothesis worked for deviations in the orbit of Uranus and were confirmed by the discovery of a new planet, Le Verrier had good reasons to think it would work in the case of Mercury as well.

The planet Vulcan had many supporters long after Le Verrier's death, but the discovery of a new way to think about spacetime and gravity put this support to rest. This required intellectual honesty, for as much as people wanted to find the planet which would explain Mercury's orbit, they had to admit there search had not been successful. It's hard to devote your life to a hypothesis that turns out to be wrong. This means it also required some intellectual humility to admit that they were wrong, and that the 50 year quest for Vulcan had been in vain.

But this does not mean the end of curiosity! Curiosity is such that we can be mistaken in what we thought was true, and use our mistake as fuel to start moving a new direction to see what we can discover. Our curiosity should never be based on being right, but on wanting to figure things out. Le Verrier was equally virtuous in his search for Vulcan as he was in his search for Neptune, even though he was right in one case and wrong in the other.

Charity

All of the aforementioned virtues are worth cultivating. But there is one more worth reminding ourselves is a virtue: charity. Recall in section 1.1 "The Principle of Charity" that we discussed the Principle of Charity. Review that for a slightly more complete discussion of the virtue of charity.

To be charitable is to attribute the best intentions and strongest justifications to someone else. To interpret a set of actions charitably is to try to see those actions in terms of the most reasonable set of motivations or intentions behind them. To interpret someone's beliefs charitably is to attribute moral innocence to them as a person as far as is possible so as to give them the strongest possible benefit of the doubt. Only when you have really good reasons for doing so might you think of someone else as irrational, vicious (in the sense meaning the opposite of virtuous), or petty. Charity, then, is a habit of interpreting actions and beliefs in a good light —a rational and moral light.

All of these dispositions have their appropriate limits, of course: many beliefs and actions are just wrongheaded or irrational or bigoted and we needn't bend ourselves in pretzel knots trying to interpret them charitably. Many of our own beliefs are things we have really good reason for believing, so we don't need to be so humble that we refuse to believe anything. Some of us, furthermore, are really in a better position to know things and to reason about them. A false sense of humility stops being honest at a certain point.

---

[1] Alfred North Whitehead, *Modes of Thought*. New York: The Macmillan Co., 1938, p. 16.

4.3: Some Intellectual Vices

A helpful way to practice these intellectual virtues is to see ways that we fail to practice them, so that we can learn from those and avoid them. Learning from how we think badly is a great way to learn how to think well. "Vice" is the traditional name for lacking in virtue. Typically vices are themselves character traits or dispositions, defects in who we are that cause us to act poorly. We're going to focus on vices that have traditionally been called "fallacies," a term that is less helpful because it suggests that such actions are *always* wrong. This is not true. With all virtues and vices, context matters for whether an action expresses virtue or vice in a particular situation. So in general, we prefer the term "vice" rather than "fallacy". However, since some of the content that follows is excerpted from more traditional textbooks, you will see the language of fallacies used, but know that we're always talking about intellectual vices, ways we fall short of being intellectually virtuous.

Most of the following vices are not themselves character traits, but expressions of character traits. Using the term "vice" more expansively is not contrary to virtues and vices being fundamentally character traits. Philosopher Quassim Cassam notes that a vice is anything in us that is "likely to impede effective and responsible inquiry" (165),[1] and this includes both bad character traits like dishonesty or arrogance, and their expression in behaviors like wishful thinking or stubbornly ignoring evidence contrary to one's beliefs. The vices below are all ways of not thinking well because of failures in our intellectual character.

Since we noted that these vices are also labeled fallacies, it's important to realize that these are not the same kinds of fallacies we will encounter later in Chapter 7 when we discuss "logical fallacies." Logical fallacies describe rules of logic in which an inference always leads to a conclusion that does not follow from the premises. Logical fallacies are always bad. The intellectual vices **are not** logical fallacies. They have to do with how we behave badly in our thinking or in our conversations with other people. This is another reason why labeling both types of mistakes "fallacies" is unhelpful, because it suggests they're both the same kinds of things. Some textbooks call the logical fallacies "formal fallacies" and the vices in this chapter "informal fallacies," to show that the first kind are mistakes of logical form and the second kind are not. We think it is most helpful to simply call them something else: intellectual vices.

It's worth noting that the presence of intellectual vices means we haven't gone about reasoning with virtue. It does **not** mean that our conclusions, the things we believe, are in fact wrong. Vice describes our *justification* for what we believe, not the *truth* of what we believe. Even a broken clock is right twice a day, and thinking in an intellectually vicious manner doesn't mean our conclusions are false. What it does mean is we have not thought well enough to be justified in believing our conclusions. Make sure you don't use the vices to ignore the possible truth of a claim. The vices are reasons to reject a particular argument, but you should always ask yourself, "What would this argument look like if it was virtuously argued?"

Okay, how will we progress from here? There are two sorts of vices we'll discuss in this chapter: **Vices of Relevance** and **Vices of Presumption**. We'll go through each vice and offer some examples. Most of the rest of this chapter is pulled from Van Cleave and Knachel. There are also fallacies in Chapter 2 and many more in Chapter 8.

Vices of Relevance

Vices of relevance are ways of making arguments or critiquing arguments that have no relevance to the arguments themselves. When we are not intellectually honest or humble, or when we lack curiosity and charity, then we tend to be more focused on winning arguments or proving our friends wrong than seeking which conclusions actually have the strongest justification. That tempts us to make arguments that are psychologically or socially successful, but not actually good arguments (because they depend on irrelevant details). Keep in mind that not every topic shift in an argument is a vice of relevance, because sometimes the new topic *is* relevant to the argument. We'll see an excellent example of this in our first vice of relevance.

Ad Hominem Attack/Argument Against the Person and Genetic Fallacy

The vice of an **Ad Hominem Attack** occurs when someone unfairly attacks the character and motives of the arguer instead of their argument. Recall that as charitable thinkers, we are trying to separate the arguer from their argument and address the latter on its own merits. If even the most loathsome person makes a good argument, that argument remains valid or strong regardless of the failings of the person making it. But not every Ad Hominem is a vice; sometimes people put forward bad arguments because of a lack of virtue. For example, if a person makes an argument, not for the sake of truth but to prove that they are smarter than others, then it is an appropriate response to avoid the argument and instead criticize their lack of honesty or curiosity. Notice that this does not render their argument bad, it simply avoids addressing the merits of the argument until the arguer is prepared to debate those merits for the right reasons.

Because intellectual virtue is an important part of thinking well, an ad hominem critique is appropriate when someone (including ourselves!) is not thinking virtuously. As Qassim Cassam writes, “The evaluation of the justificational status of a particular belief is closely related to the evaluation of the believer” (2016: 175). If we think someone has made a good argument, we’re saying they are thinking well, and this means they are thinking virtuously. Cassam elaborates, “A justified belief is characteristically one which arises through the exercise of intellectual virtue. In evaluating a belief as justified we are in effect commending the believer” (2016: 176).

Okay, so when is an ad hominem attack a vice? If we think about Cassam’s proposal above, then ad hominem is a vice whenever we attacks someone’s character (instead of their argument) for reasons *other than* their lack of intellectual virtue. If I say of someone, “I don’t think your argument is honest about the reasons against your position,” I’m fairly criticizing a lack of virtue. But if I say of someone, “I don’t think your argument is good because you’re a loan shark,” I have exemplified a vice; I have made an argument without virtue.

There are three main types of ad hominem attack:

1. Abusive: you simply attack the character or rationality of your opponent (or a group to which the opponent belongs like "liberals" or "pro-lifers")

2. Circumstantial: as above in our cartoon, you point to circumstances which make your opponent untrustworthy or suspect.

3. *Tu Quoque*: Latin meaning "you too!" You point to similar faults in your opponent when your actions or character is called into question. More generally: you point to a fault elsewhere to draw attention away from the fault being discussed.

> **From Matthew J. Van Cleave's Introduction to Logic and Critical Thinking, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.**
>
> “Ad hominem” is a Latin phrase that can be translated into English as the phrase, “against the man.” In an ad hominem fallacy, instead of responding to (or attacking) the argument a person has made, one attacks the person him or herself. In short, one attacks the person making the argument rather than the argument itself. Here is an anecdote that reveals an ad hominem fallacy (and that has actually occurred in my ethics class before).
>
> > A philosopher named Peter Singer had made an argument that it is morally wrong to spend money on luxuries for oneself rather than give all of your money that you don’t strictly need away to charity. The argument is actually an argument from analogy (whose details I discussed in section 3.3), but the essence of the argument is that there are every day in this world children who die preventable deaths, and there are charities who could save the lives of these children if they are funded by individuals from wealthy countries like our own. Since there are things that we all regularly buy that we don’t need (e.g., Starbuck’s lattes, beer, movie tickets, or extra clothes or shoes we don’t really need), if we continue to purchase those things rather than using that money to save the lives of children, then we are essentially contributing to the deaths of those children if we choose to continue to live our lifestyle of buying things we don’t need, rather than donating the money to a charity that will save lives of children in need. In response to Singer’s argument, one student in the class asked: “Does Peter Singer give his money to charity? Does he do what he says we are all morally required to do?”
>
> The implication of this student’s question (which I confirmed by following up with her) was that if Peter Singer himself doesn’t donate all his extra money to charities, then his argument isn’t any good and can be dismissed. But that would be to commit an ad hominem fallacy. Instead of responding to the argument that Singer had made, this student attacked Singer himself. That is, they wanted to know how Singer lived and whether he was a hypocrite or not. Was he the kind of person who would tell us all that we had to live a certain way but fail to live that way himself? But all of this is irrelevant to assessing Singer’s argument. Suppose that Singer didn’t donate his excess money to charity and instead spent it on luxurious things for himself. Still, the argument that Singer has given can be assessed on its own merits. Even if it were true that Peter Singer was a total hypocrite, his argument may nevertheless be rationally compelling. And it is the quality of the argument that we are interested in, not Peter Singer’s personal life and whether or not he is hypocritical. Whether Singer is or isn’t a hypocrite, is irrelevant to whether the argument he has put forward is strong or weak, valid or invalid. The argument stands on its own and it is that argument rather than Peter Singer himself that we need to assess.
>
> Nonetheless, there is something psychologically compelling about the question: Does Peter Singer practice what he preaches? I think what makes this question seem compelling is that humans are very interested in finding “cheaters” or hypocrites—those

who say one thing and then do another. Evolutionarily, our concern with cheaters makes sense because cheaters can't be trusted and it is essential for us (as a group) to be able to pick out those who can't be trusted. That said, whether or not a person giving an argument is a hypocrite is irrelevant to whether that person's argument is good or bad. So there may be psychological reasons why humans are prone to find certain kinds of ad hominem fallacies psychologically compelling, even though ad hominem fallacies are not rationally compelling.

Not every instance in which someone attacks a person's character is an ad hominem fallacy. Suppose a witness is on the stand testifying against a defendant in a court of law. When the witness is cross examined by the defense lawyer, the defense lawyer tries to go for the witness's credibility, perhaps by digging up things about the witness's past. For example, the defense lawyer may find out that the witness cheated on her taxes five years ago or that the witness failed to pay her parking tickets. The reason this isn't an ad hominem fallacy is that in this case the lawyer is trying to establish whether what the witness is saying is true or false and in order to determine that we have to know whether the witness is trustworthy. These facts about the witness's past may be relevant to determining whether we can trust the witness's word. In this case, the witness is making claims that are either true or false rather than giving an argument. In contrast, when we are assessing someone's argument, the argument stands on its own in a way the witness's testimony doesn't. In assessing an argument, we want to know whether the argument is strong or weak and we can evaluate the argument using the logical techniques surveyed in this text. In contrast, when a witness is giving testimony, they aren't trying to argue anything. Rather, they are simply making a claim about what did or didn't happen. So although it may seem that a lawyer is committing an ad hominem fallacy in bringing up things about the witness's past, these things are actually relevant to establishing the witness's credibility. In contrast, when considering an argument that has been given, we don't have to establish the arguer's credibility because we can assess the argument they have given on its own merits. The arguer's personal life is irrelevant.

Figure *[Math Processing Error]*: "Look," says the bird, "if we start walking around on the ground all the time, the cat will get us, I know I have a personal stake in this because I'd prefer not to be eaten, but my argument would stand even if I myself were a cat!" (Image Credit: Otto Speckter in *Picture Fables*)

Tu Quoque

Tu Quoque is a version of the Ad Hominem fallacy. Here's Van Cleave again.

"Tu quoque" is a Latin phrase that can be translated into English as "you too" or "you, also." The tu quoque fallacy is a way of avoiding answering a criticism by bringing up a criticism of your opponent rather than answer the criticism. For example, suppose that two political candidates, A and B, are discussing their policies and A brings up a criticism of B's policy. In response, B brings up her own criticism of A's policy rather than respond to A's criticism of her policy. B has here committed the tu quoque fallacy. The fallacy is best understood as a way of avoiding having to answer a tough criticism that one may not have a good answer to. This kind of thing happens all the time in political discourse.

Tu quoque, as I have presented it, is fallacious when the criticism one raises is simply in order to avoid having to answer a difficult objection to one's argument or view. However, there are circumstances in which a tu quoque kind of response is not fallacious. If the criticism that A brings toward B is a criticism that equally applies not only to A's position but to any position, then B is right to point this fact out. For example, suppose that A criticizes B for taking money from special interest groups. In this case, B would be totally right (and there would be no tu quoque fallacy committed) to respond that not only does A take money from special interest groups, but every political candidate running for office does. That is just a fact of life in American politics today. So A really has no criticism at all to B since everyone does what B is doing and it is in many ways unavoidable. Thus, B could (and should) respond with a "you too" rebuttal and in this case that rebuttal is not a tu quoque fallacy.

Attacking causes for belief rather than reasons for belief (Genetic Fallacy)

The vice of attacking the causes for belief, sometimes called the **Genetic Fallacy**, requires learning the difference between causes and reasons. Perhaps I trust my physician because my best friend goes to the same physician. The 'because' here is an explanation for why I trust them. But if you were to ask me why I trust my physician, I might say, "Because she is the most highly-rated general practitioner in my area." Now I have given you a reason for trusting her. Both can be true descriptions of my trust: the cause of my trust is that my best friend trusts her, and the reason I think my trust is justified is that she is so highly rated.

When criticizing an argument, we want to criticize the reasons for belief, not the causes. The genetic fallacy occurs when, for example, instead of looking at your beliefs as they stand on their own, I look at the role those beliefs play in your psychology or the psychological origins of those beliefs. I might say that you only believe in the free market because your father believes in the free market. That's not an attack against the belief itself. At best it amounts to the claim that you don't have any justification for believing it, only an explanation for how you came to believe it.

That'd be like critiquing a particular golf club because a bad brand name manufactured it. It's still a perfectly good golf club no matter who made it. We should critique the golf club on the basis of its usefulness as a golf club, not on the basis of where it was made. Note that it might be reasonable not to *trust* a bad brand when making a purchase, but if the reviews come in and it's a fine golf club, then its origin is irrelevant.

From Matthew J. Van Cleave's Introduction to Logic and Critical Thinking, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

The genetic fallacy occurs when one argues (or, more commonly, implies) that the origin of something (e.g., a theory, idea, policy, etc.) is a reason for rejecting (or accepting) it. For example, suppose that Jack is arguing that we should allow physician assisted suicide and Jill responds that that idea first was used in Nazi Germany. Jill has just committed a genetic fallacy because she is implying that because the idea is associated with Nazi Germany, there must be something wrong with the idea itself. What she should have done instead is explain what, exactly, is wrong with the idea rather than simply assuming that there must be something wrong with it since it has a negative origin. The origin of an idea has nothing inherently to do with its truth or plausibility. Suppose that Hitler constructed a mathematical proof in his early adulthood (he didn't, but just suppose). The validity of that mathematical proof stands on its own; the fact that Hitler was a horrible person has nothing to do with whether the proof is good. Likewise with any other idea: ideas must be assessed on their own merits and the origin of an idea is neither a merit nor demerit of the idea.

Although genetic fallacies are most often committed when one associates an idea with a negative origin, it can also go the other way: one can imply that because the idea has a positive origin, the idea must be true or more plausible. For example, suppose that Jill argues that the Golden Rule is a good way to live one's life because the Golden Rule originated with Jesus in the Sermon on the Mount (it didn't, actually, even though Jesus does state a version of the Golden Rule). Jill has committed the genetic fallacy in assuming that the (presumed) fact that Jesus is the origin of the Golden Rule has anything to do with whether the Golden Rule is a good idea.

I'll end with an example from William James's seminal work, The Varieties of Religious Experience. In that book (originally a set of lectures), James considers the idea that if religious experiences could be explained in terms of neurological causes, then the legitimacy of the religious experience is undermined. James, being a materialist who thinks that all mental states are physical states—ultimately a matter of complex brain chemistry, says that the fact that any religious experience has a physical cause does not undermine that veracity of that experience. Although he doesn't use the term explicitly, James claims that the claim that the physical origin of some experience undermines the veracity of that experience is a genetic fallacy. Origin is irrelevant for assessing the veracity of an experience, James thinks. In fact, he thinks that religious dogmatists who take the origin of the Bible to be the word of God are making exactly the same mistake as those who think that a physical explanation of a religious experience would undermine its veracity. We must assess ideas for their merits, James thinks, not their origins.

How do intellectually virtuous thinkers avoid making ad hominem attacks when they're inappropriate? Well, if we're intellectually honest, we will emphasize substance over motives. We will be slow to question someone's motives behind an argument, and instead start by charitably focusing on the substance of what they have to say. Of course, "slow" does not mean never, and sometimes a person's behavior and manner of argument will convince us that bad motives are a factor, but we should not *start* from a place of assuming ill-intent.

In general, we should be slow to cast aspersions on another person's character or intelligence. Just because we think they have made a bad argument does not mean we should attribute this to a lack of ability or integrity on their part. Some people (ourselves most of all!) simply make mistakes. By focusing on their argument, we continue to treat them as an equal dialogue partner, someone whose views are worthy of our curiosity and our charity. This often makes the dialogue proceed better and with more insight. Again, "slow" does not mean never, and sometimes a person who is behaving belligerently needs to be told that their conduct makes them unfit for continued dialogue. But if we resort to such "last measures," it should always be in the hope of helping a person become more intellectually virtuous so that they can rejoin the conversation, and certainly not with the secret motive of getting them to agree with us or winning the debate.

Mansplaining

Sometimes we are so confident we're right, we begin to explain why we're right in a manner and tone that is aggressive, domineering, and keeps the other person from contributing. This has become known as **Mansplaining**, a term coined because many women have experienced their ideas ignored or discredited by men who speak at them in a condescending manner. But mansplaining can be practiced by persons of any gender towards persons of any other gender. It often involves telling the other person how they feel or should feel, what they believe, and why their perspective doesn't matter.

Consider Hanuni. Hanuni shares with a friend her anxieties concerning the Russian invasion of Ukraine. She tells her friend, "I have a hard time focusing on my daily responsibilities because I feel overwhelmed at the thought of Ukrainians right now fighting and dying just to be free enough to carry out their daily responsibilities." Her friend, Hiari, replies to her, "Come'on, that's not how you feel. You don't know what it's like to be a Ukrainian, and you've never been to war, so you really have no business assuming you know what they're going through. You really should be counting your own blessings rather than worrying about things that are not your problem." Think about how Hiari's response shuts Hanuni down and makes her feel that she's wrong to care about the situation in Ukraine or to empathize with people in other situations. Most significantly, Hiari's comment erases Hanuni's voice and contribution.

A somewhat high profile instance of mansplaining occurred during the 2017 U. S. Senate debate on the confirmation of Senator Jeff Sessions (Alabama) to the office of the Attorney General of the United States. The confirmation process was contentious in part because of concerns about Senator Sessions record on civil rights. To speak to this issue, fellow Senator Elizabeth Warren (Massachusetts) reminded the Senate of former Senator Ted Kennedy's (also of Massachusetts) objections back in 1986 to Sessions being appointed to a judgeship because of concerns over suppression of black votes in his area of authority. She then proceeded to read a letter on the Senate floor authored by Coretta Scott King, the widow of civil rights leader Martin Luther King, Jr., written to the Senate Judiciary Committee in 1986 opposing Sessions' confirmation to a judgeship.

As Senator Warren was reading from King's letter, the presiding Senate Chair Steve Daines interrupted her twice to remind her that Senate rules prohibit casting aspersions on other Senators. After some back and forth, he permitted her to continue reading King's letter. However, shortly after resuming her reading, then Senate Majority Leader Mitch McConnell interrupted her, insisting that she was slandering Senator Sessions character from the floor. He called for a vote on whether she would be allowed to continue her speech, and the Senate voted to terminate her speaking time. Later in the Senate debate, another male Senator read the letter by King without objection.

Shortly after Elizabeth Warren was told to sit down, Majority Leader McConnell explained the events in the following manner, "Here is what transpired. Senator Warren was giving a lengthy speech. She had appeared to violate the rule. She was warned. She was given an explanation. Nevertheless, she persisted." McConnell's interruptions of Warren's speech, and his domineering chastisement lecturing her on why she was not permitted to continue, was more focused on explaining at her why her voice wasn't going to be included than dialoguing with her on what she had to say.

Mansplaining is never good, but it's important that we do not label everyone who criticizes what we believe as engaging in mansplaining. Intellectually virtuous people are teachable and allow others to help them see their mistakes. Sometimes people will dismiss the arguments of others as mansplaining when in fact they're only voicing disagreement. Dismissing a reasonable counter-argument as mansplaining is in fact a type of mansplaining—another way to shut down someone's voice. So it's important to correctly identify cases of mansplaining and not use the concept as a means to avoid having to listen to anyone who challenges our thinking.

Straw Argument

Figure *[Math Processing Error]*: Do you want to build a snowman? And then critique his position on global warming? (Image Credit: Otto Speckter in *Picture Fables*)

The vice of constructing a straw argument happens when someone (willfully or mistakenly) misinterprets someone else's argument or position. We also might call it creating a **Straw Argument**, **Straw Figure**, **Straw Person,** or **Straw Man.**

The opponent's argument or position is characterized *uncharitably* so as to make it seem ridiculous or indefensible. It is a failure of charity because the person is attacking *an irrelevant argument*, rather than the argument they actually gave. Imagine someone building a straw doll and fighting that instead of their actual opponent. No one would think they had won the fight.

From Matthew J. Van Cleave's Introduction to Logic and Critical Thinking, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

Suppose that my opponent has argued for a position, call it position A, and in response to his argument, I give a rationally compelling argument against position B, which is related to position A, but is much less plausible (and thus much easier to refute). What I have just done is attacked a straw man—a position that “looks like” the target position, but is actually not that position. When one attacks a straw man, one commits the straw man fallacy. The straw man fallacy misrepresents one’s opponent’s argument and is thus a kind of irrelevance. Here is an example.

> Two candidates for political office in Colorado, Tom and Fred, are having an exchange in a debate in which Tom has laid out his plan for putting more money into health care and education and Fred has laid out his plan which includes earmarking more state money for building more prisons which will create more jobs and, thus, strengthen Colorado’s economy. Fred responds to Tom’s argument that we need to increase funding to health care and education as follows: “I am surprised, Tom, that you are willing to put our state’s economic future at risk by sinking money into these programs that do not help to create jobs. You see, folks, Tom’s plan will risk sending our economy into a tailspin, risking harm to thousands of Coloradans. On the other hand, my plan supports a healthy and strong Colorado and would never bet our state’s economic security on idealistic notions that simply don’t work when the rubber meets the road.”

Fred has committed the straw man fallacy. Just because Tom wants to increase funding to health care and education does not mean he does not want to help the economy. Furthermore, increasing funding to health care and education does not entail that fewer jobs will be created.

Fred has attacked a position that is not the position that Tom holds, but is in fact a much less plausible, easier to refute position. However, it would be silly for any political candidate to run on a platform that included “harming the economy.” Presumably no political candidate would run on such a platform. Nonetheless, this exact kind of straw man is ubiquitous in political discourse in our country.

Here is another example.

✔ Example *[Math Processing Error]*

Nancy has just argued that we should provide middle schoolers with sex education classes, including how to use contraceptives so that they can practice safe sex should they end up in the situation where they are having sex. Fran responds: “proponents of sex education try to encourage our children to a sex-with-no-strings-attached mentality, which is harmful to our children and to our society.”

Fran has committed the straw man (or straw woman) fallacy by misrepresenting Nancy’s position. Nancy’s position is not that we should encourage children to have sex, but that we should make sure that they are fully informed about sex so that if they

> do have sex, they go into it at least a little less blindly and are able to make better decision regarding sex.
>
> As with other fallacies of relevance, straw man fallacies can be compelling on some level, even though they are irrelevant. It may be that part of the reason we are taken in by straw man fallacies is that humans are prone to "demonize" the "other"—including those who hold a moral or political position different from our own. It is easy to think bad things about those with whom we do not regularly interact. And it is easy to forget that people who are different than us are still people just like us in all the important respects. Many years ago, atheists were commonly thought of as highly immoral people and stories about the horrible things that atheists did in secret circulated widely. People believed that these strange "others" were capable of the most horrible savagery. After all, they may have reasoned, if you don't believe there is a God holding us accountable, why be moral? The Jewish philosopher, Baruch Spinoza, was an atheist who lived in the Netherlands in the 17th century. He was accused of all sorts of things that were commonly believed about atheists. But he was in fact as upstanding and moral as any person you could imagine. The people who knew Spinoza knew better, but how could so many people be so wrong about Spinoza? I suspect that part of the reason is that since at that time there were very few atheists (or at least very few people actually admitted to it), very few people ever knowingly encountered an atheist. Because of this, the stories about atheists could proliferate without being put in check by the facts. I suspect the same kind of phenomenon explains why certain kinds of straw man fallacies proliferate. If you are a conservative and mostly only interact with other conservatives, you might be prone to holding lots of false beliefs about liberals. And so maybe you are less prone to notice straw man fallacies targeted at liberals because the false beliefs you hold about them incline you to see the straw man fallacies as true.

Thinking with virtue means that when others explicitly deny a view, we should be slow to attribute this view to them. This does not mean we never do so; again, if someone is acting in bad faith and we think they are pretending to hold a view different than the one they assert, we might need to make clear their hidden agenda. But notice this is no longer a straw argument, if we're right in our suspicion. Nonetheless, we start from a place of being slow to do this, wanting to take people at face value first before assuming they don't believe what they are claiming.

A related practice in virtue is to be slow to attribute to others views that are clearly false, implausible, or lie at the extremes of human belief. Again, sometimes we have to do this because there are people who believe false, implausible, or extremist views. But we start from a place of charitably assuming rationality and truth in people, being slow to change our assumption.

A really useful way to assist with this is to summarize the other person's views and arguments back to them before making a critique. If we stop ourselves and explain to someone else what we think they are arguing, it (a) gives them an opportunity to clarify first before we make objections, and (b) it shows them we are acting in good faith and that they can trust us to not construct straw arguments out of what they said.

Red Herring

Figure *[Math Processing Error]*: Even the goodest boiz get distracted easily. SQUIRREL! (Image Credit: Otto Speckter in *Picture Fables*)

A herring is a pungent fish, especially in the days before refrigeration. William Cobbett claimed to have used this as a boy to lure unsuspecting hounds and their unsuspecting hunters away from their intended prey. Cobbett wanted the rabbit for himself, so he drug a herring on the ground to make a stench trail, drawing the hound away from the rabbit's hole.

Interesting trick! But what does this have to do with reasoning well? Simple: one way that people reason improperly is by **not staying on topic.** If you start talking about one thing, but end up talking about another thing, chances are either you or your conversant have committed the vice of a red herring. This is where you intentionally or unintentionally change the subject. Often it happens when a politician doesn't want to answer a question. "I don't want to talk about jobs, I want to talk about the brave men and women who serve in our nation's proud military…" It's a great way to get around having to answer a question.

A Red Herring is sometimes hard to distinguish from a Straw Figure. Let's focus on the key difference for one second. In a straw figure, the offender is attacking an *irrelevant argument* instead of the actual argument of their opponent. In a red herring, the offender is introducing an *irrelevant topic* and discussing that instead of the topic at hand. We don't change topics in a straw figure, we just start talking about a different argument *on the same topic*.

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

A fictional example can illustrate the technique. Consider Frank, who, after a hard day at work, heads to the tavern to unwind. He has far too much to drink, and, unwisely, decides to drive home. Well, he's swerving all over the road, and he gets pulled over by the police. Let's suppose that Frank has been pulled over in a posh suburb where there's not a lot of crime. When the police officer tells him he's going to be arrested for drunk driving, Frank becomes belligerent:

> "Where do you get off? You're barely even real cops out here in the 'burbs. All you do is sit around all day and pull people over for speeding and stuff. Why don't you go investigate some real crimes? There's probably some unsolved murders in the inner city they could use some help with. Why do you have to bother a hard-working citizen like me who just wants to go home and go to bed?"

Frank is committing the red herring fallacy (and not very subtly). The issue at hand is whether or not he deserves to be arrested for driving drunk. He clearly does. Frank is not comfortable arguing against that position on the merits. So he changes the subject—to one about which he feels like he can score some debating points. He talks about the police out here in the suburbs, who, not having much serious crime to deal with, spend most of their time issuing traffic violations. Yes, maybe that's not as taxing a job as policing in the city. Sure, there are lots of serious crimes in other jurisdictions that go unsolved. But that's beside the point! It's a distraction from the real issue of whether Frank should get a DUI.

Politicians use the red herring fallacy all the time. Consider a debate about Social Security—a retirement stipend paid to all workers by the federal government. Suppose a politician makes the following argument:

> We need to cut Social Security benefits, raise the retirement age, or both. As the baby boom generation reaches retirement age, the amount of money set aside for their benefits will not be enough cover them while ensuring the same standard of living for future generations when they retire. The status quo will put enormous strains on the federal budget going forward, and we are already dealing with large, economically dangerous budget deficits now. We must reform Social Security.

Now imagine an opponent of the proposed reforms offering the following reply:

> Social Security is a sacred trust, instituted during the Great Depression by FDR to insure that no hard-working American would have to spend their retirement years in poverty. I stand by that principle. Every citizen deserves a dignified retirement. Social Security is a more important part of that than ever these days, since the downturn in the stock market has left many retirees with very little investment income to supplement government support.

The second speaker makes some good points, but notice that they do not speak to the assertion made by the first: Social Security is economically unsustainable in its current form. It's possible to address that point head on, either by making the case that in fact the economic problems are exaggerated or non-existent, or by making the case that a tax increase could fix the problems. The respondent does neither of those things, though; he changes the subject, and talks about the importance of dignity in retirement. I'm sure he's more comfortable talking about that subject than the economic questions raised by the first speaker, but it's a distraction from that issue—a red herring.

Perhaps the most blatant kind of red herring is evasive: used especially by politicians, this is the refusal to answer a direct question by changing the subject. Examples are almost too numerous to cite; to some degree, no politician ever answers difficult questions straightforwardly (there's an old axiom in politics, put nicely by Robert McNamara: "Never answer the question that is asked of you. Answer the question that you wish had been asked of you.").

A particularly egregious example of this occurred in 2009 on CNN's Larry King Live. Michele Bachmann, Republican Congresswoman from Minnesota, was the guest. The topic was "birtherism," the (false) belief among some that Barack Obama was not in fact born in America and was therefore not constitutionally eligible for the presidency. After playing a clip of Senator Lindsey Graham (R, South Carolina) denouncing the myth and those who spread it, King asked Bachmann whether she agreed with Senator Graham. She responded thus:

> "You know, it's so interesting, this whole birther issue hasn't even been one that's ever been brought up to me by my constituents. They continually ask me, where's the jobs? That's what they want to know, where are the jobs?"

Bachmann doesn't want to respond directly to the question. If she outright declares that the "birthers" are right, she looks crazy for endorsing a clearly false belief. But if she denounces them, she alienates a lot of her potential voters who believe the falsehood. Tough bind. So she blatantly, and rather desperately, tries to change the subject. Jobs! Let's talk about those instead. Please?

Irrelevant Appeals

Any kind of appeal to a factor, consideration, or reason that isn't relevant to the argument at hand (but is used *as a reason* rather than as a mere distraction—A Red Herring is a distraction, not an irrelevant reason) is called an Irrelevant Appeal. The premises aren't relevant to the truth or falsity of the conclusion because whether or not the conclusion is true doesn't depend at all on whether or not the premises are true.

The core Irrelevant Appeals to Know:

- Appeal to Unqualified/False Authority
- Appeal to Force
- Appeal to Popularity/to the People/Bandwagon
- Appeal to Consequences

Appeal to Unqualified Authority

Note that this is sometimes called the "Appeal to Authority", but we trust authorities all the time about lots of things and we're right to do so. The fallacy is when we trust an authority on one subject (or perhaps someone who is not an authority on anything at all) to speak on another subject.

Figure *[Math Processing Error]*: No matter the fact that you're my elder, Mr. Turkey, you're no expert on Quantum Physics! (Image Credit: Otto Speckter in *Picture Fables*)

From Matthew J. Van Cleave's *Introduction to Logic and Critical Thinking*, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

In a society like ours, we have to rely on authorities to get on in life. For example, the things I believe about electrons are not things that I have ever verified for myself. Rather, I have to rely on the testimony and authority of physicists to tell me what electrons are like. Likewise, when there is something wrong with my car, I have to rely on a mechanic (since I lack that expertise) to tell me what is wrong with it. Such is modern life. So there is nothing wrong with needing to rely on authority figures in certain fields (people with the relevant expertise in that field)—it is inescapable. The problem comes when we invoke someone whose expertise is not relevant to the issue for which we are invoking it. For example, suppose that a group of doctors sign a petition to prohibit abortions, claiming that abortions are morally wrong. If Bob cites that fact that these doctors are against abortion, therefore abortion must be morally wrong, then Bob has committed the appeal to authority fallacy. The problem is that doctors are not authorities on what is morally right or wrong. Even if they are authorities on how the body works and how to perform certain procedures (such as abortion), it doesn't follow that they are authorities on whether or not these procedures should be performed—the ethical status of these procedures. It would be just as much an appeal to consequences fallacy if Melissa were to argue that since some other group of doctors supported abortion, that shows that it must be morally acceptable. In either case, since doctors are not authorities on moral issues, their opinions on a moral issue like abortion is irrelevant. In general, an appeal to authority fallacy occurs when someone takes what an individual says as evidence for some claim, when that individual has no particular expertise in the relevant domain (even if they do have expertise in some other, unrelated, domain).

Appeal to Force

An appeal to force is an irrelevant appeal because it apparently argues that some proposition is true, but uses as justification for that claim a threat on the listener. If you don't believe this, then you will suffer bad consequences. But that's not a reason to believe the proposition. That's a reason to make yourself believe it or to act as if you believe it. A good argument actually gives you reason to believe the conclusion and an appeal to force does no such thing!

The following is from: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

Perhaps the least subtle of the fallacies is the appeal to force, in which you attempt to convince your interlocutor to believe something by threatening him. Threats pretty clearly distract one from the business of dispassionately appraising premises' support for conclusions, so it's natural to classify this technique as a Fallacy of Distraction.

There are many examples of this technique throughout history. In totalitarian regimes, there are often severe consequences for those who don't toe the party line (see George Orwell's 1984 for a vivid, though fictional, depiction of the phenomenon). The Catholic Church used this technique during the infamous Spanish Inquisition: the goal was to get non-believers to accept Christianity; the method was to torture them until they did.

An example from much more recent history: when it became clear in 2016 that Donald Trump would be the Republican nominee for president, despite the fact that many rank-and-file Republicans thought he would be a disaster, the Chairman of the Republican National Committee (allegedly) sent a message to staffers informing them that they could either support Trump or leave their jobs. Not a threat of physical force, but a threat of being fired; same technique.

Again, the appeal to force is not usually subtle. But there is a very common, very effective debating technique that belongs under this heading, one that is a bit less overt than explicitly threatening someone who fails to share your opinions. It involves the sub-conscious, rather than conscious, perception of a threat.

Here's what you do: during the course of a debate, make yourself physically imposing; sit up in your chair, move closer to your opponent, use hand gestures, like pointing right in their face; cut them off in the middle of a sentence, shout them down, be angry and combative. If you do these things, you're likely to make your opponent very uncomfortable—physically and emotionally. They might start sweating a bit; their heart may beat a little faster. They'll get flustered and maybe trip over their words. They may lose their train of thought; winning points they may have made in the debate will come out wrong or not at all. You'll look like the more effective debater, and the audience's perception will be that you made the better argument.

But you didn't. You came off better because your opponent was uncomfortable. The discomfort was not caused by an actual threat of violence; on a conscious level, they never believed you were going to attack them physically. But you behaved in a way that triggered, at the sub-conscious level, the types of physical/emotional reactions that occur in the presence of an actual physical threat. This is the more subtle version of the appeal to force. It's very effective and quite common (watch cable news talk shows and you'll see it; Bill O'Reilly is the master).

Ad Populum

Figure *[Math Processing Error]*: I don't care how popular bear jousting is, it's just wrong! (Image Credit: Otto Speckter in *Picture Fables*)

Appeal to the People, to Popularity, Nose-Counting Fallacy, Bandwagon Fallacy, *argumentum ad populum* are all names for the same thing: appealing to the popularity of a thing or idea or practice in order to justify that thing or idea or practice. In an argument, one appeals to the popularity of a conclusion and then uses that popularity as a basis for inferring that the conclusion is true.

The popularity of a new smartphone or computer might be used to justify it's status as the best available. The popularity of a politician might be used to justify the claim that they should be President. The popularity of a person might be used to attempt to exonerate them from a crime or protect them from criticism. In each case, mere popularity doesn't mean we should believe something is good or worthy of special consideration.

The popularity of belief in God might be used as evidence that God exists. After all, that many people can't be wrong, right? Alternatively, the popularity among scientists of belief in an atheistic universe might be used as evidence that God doesn't exist. After all, that many scientists can't be wrong, can they?

In reality, the popularity of a belief doesn't give us reason to think that belief is true. After all, there have been lots of popular ideas in the past that turned out to be not only false, but morally abhorrent!

Appeal to Consequences

Appeal to consequences is yet another "irrelevant appeal" vice. Again something which isn't relevant to the truth or falsity of the conclusion is appealed to in arguing for that conclusion. It won't help though, since it's not relevant!

From Matthew J. Van Cleave's *Introduction to Logic and Critical Thinking*, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

The appeal to consequences fallacy is like the reverse of the genetic fallacy: whereas the genetic fallacy consists in the mistake of trying to assess the truth or reasonableness of an idea based on the origin of the idea, the appeal to consequences fallacy consists in the mistake of trying to assess the truth or reasonableness of an idea based on the (typically negative) consequences of accepting that idea. For example, suppose that the results of a study revealed that there are IQ differences between different races (this is a fictitious example, there is no such study that I know of). In debating the results of this study, one researcher claims that if we were to accept these results, it would lead to increased racism in our society, which is not tolerable. Therefore, these results must not be right since if they were accepted, it would lead to increased racism. The researcher who responded in this way has committed the appeal to consequences fallacy. Again, we must assess the study on its own merits. If there is something wrong with the study, some flaw in its design, for example, then that would be a relevant criticism of the study. However, the fact that the results of the study, if widely circulated, would have a negative effect on society is not a reason for rejecting these results as false. The consequences of some idea (good or bad) are irrelevant to the truth or reasonableness of that idea.

Notice that the researchers, being convinced of the negative consequences of the study on society, might rationally choose not to publish the study (for fear of the negative consequences). This is totally fine and is not a fallacy. The fallacy consists not in choosing not to publish something that could have adverse consequences, but in claiming that the results themselves are undermined by the negative consequences they could have. The fact is, sometimes truth can have negative consequences and falsehoods can have positive consequences. This just goes to show that the consequences of an idea are irrelevant to the truth or reasonableness of an idea.

The Fallacy Fallacy

Perhaps the most important vice to be aware of goes by the name: the Fallacy Fallacy! Remember that most other textbooks call these vices "fallacies" and remember that at the beginning of the chapter we said that whether or not one's opponent argues virtuously is irrelevant to whether or not one's opponent is in fact correct in their conclusion. They might believe the right thing for wrong reasons or they might have good reasons that just don't come through clearly when they try to explain their beliefs. Here's an example of the fallacy fallacy:

Example [Math Processing Error]

***Person E**: My opponent has argued that we should lower taxes because it would stimulate commerce. I think we should be focusing on the war we've been fighting at great cost instead of arguing about whether or not lower taxes would stimulate the economy.*

***Person F**: Well clearly my opponent has never taken a logic and critical thinking class, because they have just committed a grievous sin against reasoning: the red herring fallacy. I, therefore, conclude that we should lower taxes.*

Person E is indeed guilty of a red herring: they changed the subject to something irrelevant to the original topic. We started talking about an inference from "lowering taxes would stimulate the economy" to "we should lower taxes." But by the end of Person E's speech, we were talking about something different: a costly war our nation is fighting. The topic has changed.

That being true, though, doesn't mean that Person E is *wrong* about their conclusion. If Person E wants to cut spending on wars or raise taxes to pay for them, their reasoning badly in one particular instance does not mean that their position is wrong. It may well be that we should raise taxes. Person E just isn't the best representative of the view. Person F doesn't get my vote either, though, because they don't understand a basic truth of reasoning: just because an argument for a position is bad, doesn't mean that position is wrongheaded or incorrect.

The Fallacy Fallacy happens when someone uses the fact that a fallacy was committed to justify rejecting the *conclusion* of the fallacious argument. Avoid this sort of thinking. The fallacy fallacy might count as a vice of relevance, so we'll include it in that category for our purposes here.

Vices of Presumption

The vices in the previous section were all various examples of failing to make arguments that are relevant to the topic or argument at hand. The vices in this section have a similar unifying theme, in which something is being presumed in the premises that allows the conclusion to be inferred. That something—the presumption of the argument—is in each case not warranted. If we sneak in an assumption without actually justifying that assumption, then we're creating the illusion that we've given good reasons for what we believe, when in fact we have only presumed what we believe. Try not to presume!

Vices of presumption are all shortfalls in thinking which problematically presume their conclusion to be true in the set up or assumptions of the argument. A funny example of presumption (my classmates did this as a joke when I was in elementary school): the **complex question**. For instance, you could ask "does your mom know that you do drugs?" You would be presuming that the recipient of the question does drugs because you're only asking about their mother's knowledge. Other examples are "when are you going to stop stealing my food?" and "how do you justify to yourself that you lie to everyone all the time?". In each case, facts are being presumed that have not been agreed on as facts! This helps us get a sense of what presumption is and why it might be a problem.

Inequity in Evaluating Evidence

Consider someone who thinks whole milk ice cream is superior to frozen yogurt. Whenever someone presents evidence of the health benefits or excellent flavor in frozen yogurt, they scrutinize the evidence with great skepticism, looking for every little reason to reject the evidence. They demand near scientific thresholds to make the case for frozen yogurt. But when it comes to evidence for their own love of whole milk ice cream, they are willing to accept even anecdotal testimony or hasty statistics as bolstering their argument. What has gone wrong in this situation?

The ice cream lover is someone who applies one standard of evidence to evidence against their position, and another standard of evidence to evidence that favors their position. This is a way of presuming one is right before the evidence has been heard, such that the evaluation of the evidence serves to make sure the "right" conclusion results.

This is not how an intellectually honest and humble thinker approaches matters. They want the truth, even if it requires admitting their mistake. Thinkers disposed to virtue will be even-handed when assessing evidence, especially evidence supporting views different from their own. They will not favor evidence that supports their belief simply *because* it supports their belief, nor will they discount evidence that undermines their belief simply *because* it undermines it.

Inequity in evaluating evidence is typically is an expression of a deeper character vice in humans, **confirmation bias**. We will learn more about confirmation bias in Chapter 8.1 "Confirmation Bias". Confirmation bias is a psychological handicap in humans that once we believe something, it is easier for us to keep believing it rather than change our minds. Thus we evaluate evidence unequally because our brains are predisposed to hold on to what we already believe rather than give credence to possibilities that would require us to change our minds.

Also in Chapter 8.5 "Texas Sharpshooter" we'll learn about a fallacy of inductive reasoning nicknamed after a tall tale about a **Texas sharpshooter**. This fallacy is related to inequity in evaluating evidence, but the two vices are subtly different. Inequity in evaluating evidence is primarily about how we presume the evidence should be judged—evidence against us should be judged more stringently, while evidence in our favor should be judged more leniently. As we'll see when we learn about the Texas sharpshooter fallacy, that vice more describes a pattern of (fallacious) inductive reasoning in which we start from our conclusion and select evidence that supports it (rather than virtuous induction, where we start from our evidence and infer a conclusion). A virtuous thinker allows new evidence to dictate how they understand what conclusion is the most reasonable one. But you should see all these vices as a family: they're different ways of not thinking well about evidence. They are also different ways of not displaying the virtues of curiosity and honesty.

False Dilemma/Black and White

From Matthew J. Van Cleave's *Introduction to Logic and Critical Thinking*, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

Suppose I were to argue as follows:

Raising taxes on the wealthy will either hurt the economy or it will help it. But it won't help the economy. Therefore it will hurt the economy.

The standard form of this argument is:

1. Either raising taxes on the wealthy will hurt the economy or it will help it.

2. Raising taxes on the wealthy won't help the economy.

3. Therefore, raising taxes on the wealthy will hurt the economy.

This argument contains a fallacy called a "false dichotomy." A false dichotomy is simply a disjunction that does not exhaust all of the possible options. In this case, the problematic disjunction is the first premise: either raising the taxes on the wealthy will hurt the economy or it will help it. But these aren't the only options. Another option is that raising taxes on the wealthy will have no effect on the economy. Notice that the argument above has the form of a disjunctive syllogism:

[Math Processing Error]

However, since the first premise presents two options as if they were the only two options, when in fact they aren't, the first premise is false and the argument fails. Notice that the form of the argument is perfectly good—the argument is valid. The problem is that this argument isn't sound because the first premise of the argument commits the false dichotomy fallacy. False dichotomies are commonly encountered in the context of a disjunctive syllogism or constructive dilemma (see chapter 2).

In a speech made on April 5, 2004, President Bush made the following remarks about the causes of the Iraq war:

> Saddam Hussein once again defied the demands of the world. And so I had a choice: Do I take the word of a madman, do I trust a person who had used weapons of mass destruction on his own people, plus people in the neighborhood, or do I take the steps necessary to defend the country? Given that choice, I will defend America every time.

The false dichotomy here is the claim that:

Either I trust the word of a madman or I defend America (by going to war against Saddam Hussein's regime).

The problem is that these aren't the only options. Other options include ongoing diplomacy and economic sanctions. Thus, even if it true that Bush shouldn't have trusted the word of Hussein, it doesn't follow that the only other option is going to war against Hussein's regime. (Furthermore, it isn't clear in what sense this was needed to defend America.) That is a false dichotomy.

As with all the previous informal fallacies we've considered, the false dichotomy fallacy requires an understanding of the concepts involved. Thus, we have to use our understanding of world in order to assess whether a false dichotomy fallacy is being committed or not.

Begging the Question

From Matthew J. Van Cleave's *Introduction to Logic and Critical Thinking*, version 1.4, pp. 189-195. Creative Commons Attribution 4.0 International License.

Consider the following argument:

Capital punishment is justified for crimes such as rape and murder because it is quite legitimate and appropriate for the state to put to death someone who has committed such heinous and inhuman acts.

The premise indicator, "because" denotes the premise and (derivatively) the conclusion of this argument. In standard form, the argument is this:

1. It is legitimate and appropriate for the state to put to death someone who commits rape or murder.
2. Therefore, capital punishment is justified for crimes such as rape and murder.

You should notice something peculiar about this argument: the premise is essentially the same claim as the conclusion. The only difference is that the premise spells out what capital punishment means (the state putting criminals to death) whereas the conclusion just refers to capital punishment by name, and the premise uses terms like "legitimate" and "appropriate" whereas the conclusion uses the related term, "justified." But these differences don't add up to any real differences in meaning. Thus, the premise is essentially saying the same thing as the conclusion. This is a problem: we want our premise to provide a reason for accepting the conclusion. But if the premise is the same claim as the conclusion, then it can't possibly provide a reason for accepting the conclusion! Begging the question occurs when one (either explicitly or implicitly) assumes the truth of the conclusion in one or more of the premises. Begging the question is thus a kind of circular reasoning.

One interesting feature of this fallacy is that formally there is nothing wrong with arguments of this form. Here is what I mean. Consider an argument that explicitly commits the fallacy of begging the question. For example,

1. Capital punishment is morally permissible
2. Therefore, capital punishment is morally permissible

Now, apply any method of assessing validity to this argument and you will see that it is valid by any method. If we use the informal test (by trying to imagine that the premises are true while the conclusion is false), then the argument passes the test, since any time the premise is true, the conclusion will have to be true as well (since it is the exact same statement). Likewise, the argument is valid by our formal test of validity, truth tables. But while this argument is technically valid, it is still a really bad argument. Why? Because the point of giving an argument in the first place is to provide some reason for thinking the conclusion is true for those who don't already accept the conclusion. But if one doesn't already accept the conclusion, then simply restating the conclusion in a different way isn't going to convince them. Rather, a good argument will provide some reason for accepting the conclusion that is sufficiently independent of that conclusion itself. Begging the question utterly fails to do this and this is why it counts as an informal fallacy. What is interesting about begging the question is that there is absolutely nothing wrong with the argument formally.

Figure *[Math Processing Error]*: C'mon dog, you should trust me, my friend Rosco will tell you I'm trustworthy. I can vouch for Rosco. He's a good guy. (Image Credit: Otto Speckter in *Picture Fables*)

Whether or not an argument begs the question is not always an easy matter to sort out. As with all informal fallacies, detecting it requires a careful understanding of the meaning of the statements involved in the argument. Here is an example of an argument where it is not as clear whether there is a fallacy of begging the question:

Christian belief is warranted because according to Christianity there exists a being called "the Holy Spirit" which reliably guides Christians towards the truth regarding the central claims of Christianity.1

One might think that there is a kind of circularity (or begging the question) involved in this argument since the argument appears to assume the truth of Christianity in justifying the claim that Christianity is true. But whether or not this argument really does beg the question is something on which there is much debate within the sub-field of philosophy called epistemology ("study of knowledge"). The philosopher Alvin Plantinga argues persuasively that the argument does not beg the question, but being able to assess that argument takes patient years of study in the field of epistemology (not to mention a careful engagement with Plantinga's work). As this example illustrates, the issue of whether an argument begs the question requires us to draw on our general knowledge of the world. This is the mark of an informal, rather than formal, fallacy.

Burden of Proof Shifting

Sometimes we have a responsibility to offer evidence or proof for a claim we believe in. If I believe in dragons, then most people would think I'm responsible for proving that they exist if I expect anyone else to join me in believing in them.

Alternatively, if I believe that drivers must obey the rules of the road, most people wouldn't think I'd have to offer any justification for that belief if I brought it up in normal conversation.

Sometimes we have the **burden of proof**, but other times we do not. Here's a conversation:

Aisha: I think an alien spacecraft came and kidnapped my dog last night.

Rashid: What makes you think that?

Aisha: Well, can you prove that they didn't?

Figure *[Math Processing Error]*: (Image Credit: Otto Speckter in Picture Fables)

Something has gone wrong here, right? Aisha is making a sort of mistake: she's making an outlandish claim, but refuses to defend it or offer evidence or reasons for believing it. The vice of **Burden Shifting** is when one decides that someone else must prove them wrong when in reality they are the person with the burden of proof: one should prove oneself right!

As a general rule, whenever someone makes a positive claim about the world (like aliens kidnapped my dog), they should offer evidence or reason for believing that claim. When one makes a negative claim (like aliens *didn't* kidnap your dog), it most of the time doesn't feel like they're in the same position. It seems like they don't have to prove the negative claim unless there's already some good reason to believe in the positive claim.

This rule isn't perfect, since sometimes a belief is so commonsense that it need not be proved, but it seems to be a good general norm for where the burden of proof lies.

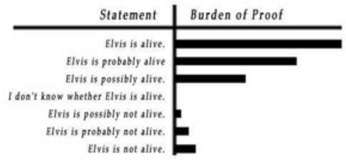


Figure *[Math Processing Error]*: (Credit: Phil Stilwell CC-License)

Alternatively, as a general rule the least plausible claim has the highest burden of proof. Since plausibility of a claim depends on all of our other beliefs, though, this is hard to adjudicate sometimes. That is fancy speak for the following idea: whoever is making the wilder claim or the claim that we're less likely to believe right away is the one with the burden of proof. This is a matter, though, of the norms of the culture we live in. In a racist society, egalitarian ideals are the ones which are "less plausible" to the elites, so they would demand more proof from someone making a claim that to us is obviously correct: that human beings are essentially equal regardless of their race. This presents a bit of a problem for those who want to use "plausibility" to decide who has the burden of proof. It suffices to say, for now, that this is simply complex and difficult to figure out.

---

[1] Quassim Cassam, Vice Epistemology, *The Monist*, Vol. 99, No. 2, Virtues (April, 2016), pp. 159-180

4.E: Chapter Four (Exercises)

Exercise [Math Processing Error]: Vices of Relevance

Identify the vice of relevance being illustrated by each argument. Remember that "no vice present" is always an option. Try to decide whether this is virtuous reasoning or not.

A. You've attack me instead of my argument, so clearly you don't know how to reason and I'm correct after all.

B. You've claimed that it's wrong to use animals and so we should all become vegans, so you're claiming that all living things are things we can't eat? Does that mean that we can't eat plants even? We have to eat *something*!!!

C. She's from Kentucky, so clearly she doesn't know a danged thing about sailing!

D. Look, you might as well just admit you're wrong. Everyone will shun you if you don't.

E. My opponent has argued that the death penalty is costly and so should be abolished, but she also supports cutting taxes! We can't cut taxes in the middle of a budget crisis.

F. You've mentioned before that you reject the tenets of capitalism, but you went to a public school, so you're not exactly an impartial judge of whether or not socialism is a good thing!

Exercise [Math Processing Error]: Vices of Presumption

Identify the vice of presumption being illustrated by each argument. Remember that "no vice present" is always an option. Try to decide whether this is virtuous reasoning or not.

A. I believe that the government is poisoning us through breakfast cereals. If you want me to eat that, you're going to have to prove to me that it's safe.

B. There's either no reason to go to space, or we should put billions into technologies which allow us to go into space. All or nothing.

C. You're either a Raiders fan or you're not a Raiders fan. Those are the only two options.

D. Look if it's bad to steal things, then it's wrong to take food that doesn't belong to you. It is bad to steal things, so it follows that you shouldn't take food from that vendor at the market.

E. The Republicans haven't championed a single non-cynical or moral policy in decades. I invite you to come up with a single example.

F. We need to bolster our space travel infrastructure, because we need to have easy and cheap access to space in the next forty years. Look, there's going to be an increased need for space travel in the near future, so we'll need cheaper access to space. A space elevator would fit the bill, and we should build one since we need to have more robust space travel infrastructure.

G. Nobody likes you. I asked everyone on the playground and not a single person said they wanted to be friends with you.

H. There will always be income inequality since there will always be rich and poor no matter what we do.

I. We shouldn't invade Iran since we shouldn't pre-emptively attack a relatively non-violent sovereign nation.

Exercise [Math Processing Error]: General Vices

Try to decide whether this is virtuous reasoning or not. If not, try to diagnose what specifically is going wrong in your own words. Then, identify the vice illustrated by each argument (can be vices of relevance or presumption). Remember that "no vice present" is always an option—it could be an example of basically virtuous reasoning!

A. You can't be a half-hearted vegetarian. You have to choose sides: either you're a vegan and an abolitionist or you're a murderer and an enslaver.

B. Eating meat is wrong because it's wrong to consume the flesh of another sentient (feeling, experiencing) being.

C. Written on a park table in Portland: "My bus costs $2.50. Does that mean I own it now?"

D. I saw some young folks at the park yesterday and they seemed to be on drugs. Isn't it terrible what is happening to our youth these days?

E. I'm pretty sure we shouldn't go to war, so that evidence that Assad is using nerve gas against his own citizens must be met with extreme suspicion.

F. We have the lowest prices since we always have lower prices than our competitors. You can be sure we always have lower prices than our competitors because we have the lowest prices available.

G. Andrew Lavin is the best textbook author because he wrote the best textbook and the author of the best textbook must be the best textbook author.

H. I won't be manipulated into believing that Area 51 isn't a storage facility for alien artifacts and specimens, you'll have to *prove* it to me using evidence and reasons.

I. I wouldn't want you to lose the next election, and I would know how to make that happen, so I expect you'll be agreeing with our policy proposal.

J. I want to go to North Korea on vacation. You'll have to prove to me it's a bad idea if you don't want me to go.

K. You want to watch the new Transformers movie? You know Michael Bay directed it, right? It's going to be terrible.

L. That cheese comes from Turkey, where they don't require pasteurization. I wouldn't recommend eating it while pregnant since listeria and other bacterial infections can be deadly to a developing fetus.

M. I understand you're frustrated with my habits, but you have some bad habits too, you know?

N. I understand that you don't want me to go on this vacation, and I respect that, but remember when you went on that vacation to visit your nephew last summer? That was a good time, right? I'm so glad you got to go on that vacation. Good times.

O. That's a slippery slope. I don't think your position can possibly be correct with reasoning like that behind it!

P. I understand you have a history of mental illness, so tell me how are we to trust your reasoning when you argue based on evidence and reasons that the Democratic Party is hopelessly corrupt and must be dissolved?

Q. I don't know. Lots of people seem pretty convinced that marriage is a love-based bond between two consenting adults, so it seems like that's what marriage is.

R. Bieber can't be the best musician. He's from Canada! They don't make good music in Canada.

S. Alanis Morrisette didn't understand the concept of irony when she wrote "Ironic". She's clearly not the most astute student of the linguistic arts.

T. That car won't run well. It was built in Russia. Cars from Russia don't tend to run well.

U. Which color do you want your car to be? Black or Gray?

V. If everyone starts believing in the tooth fairy, we'll have folks ripping out their teeth for money, so we can't encourage people to start believing in the tooth fairy.

W. Miley Cyrus said that D'Addario strings are the best guitar strings. She's a famous guitar player and musician, so I supposed D'Addario strings are really the best.

X. Rambo wasn't the greatest movie of all time. Did you know that Sylvester Stallone had a role in creating the characters and story for Creed? It was Ryan Coogler's break out film and he later went on to direct Black Panther.

Y. Veronica: I think I saw something out of the corner of my eye right now that may have been a ghost.

Hypatia: Are you saying there was definitely a ghost over there? Do you have any idea how implausible that is?

Z. Franz: There may be some reason to suspect that the threat from global warming has been overblown.

Valeria: Are you kidding me? You're a climate denier? All of the evidence points to the fact that humans have played the decisive role in warming the global climate. I can't believe you'd deny that!

AA. You've seen a ghost? That's pretty spooky. But you take anti-depressants, right? So I guess you're not that reliable [note: there is no known connection between anti-depressants and hallucinations].

BB. The CEO of Exxon Mobil has recently admitted that because of the overwhelming consensus among climate experts, we have to admit that global warming is real. But obviously they're not an impartial person so we can reject their position. They're probably doing this just for good press.

CC. My biology teacher says global warming was caused by humans burning fossil fuels and their deforestation practices. She's a scientist, so she must be right about this.

CHAPTER OVERVIEW

5: What is Logic?

Deductive logic can be used to solve logical puzzles, to understand how inferences and arguments hang together, and—most practically—to build computers. Without deductive logic, your phone would just be an expensive brick.

What is deductive logic, though?

Well, it's a method of **expression** and **analysis** of the logical form of statements, groups of statements, and inferences.

Expression: Logic is a *language* in which we can express logical relations between basic elements—we can express or represent logical structures. We do this by *abstracting away* from the English or ordinary language sentence to a statement that **only includes the logical relations**. We, in other words, take stuff away until we've only got a few basic elements and then we're ready to manipulate those basic elements in order to analyze the statements and arguments.

Analysis: Once we've got our statement, group of statements, or inference translated into a logical language, we can then use the math-like methods of analysis we'll learn to figure out:

a. What does this statement **entail** or imply? Is this statement necessarily true? Is it self-contradictory?

b. Is this set of beliefs or statements ***consistent***? Could you believe them all at once?

c. Is this argument **valid**? Do the premises *entail* the conclusion?

Here's a example:

> **Example 5.1**
>
> *If you want a ride to school, you'll need to be ready by 8.*

Translates to:

$$\textit{(Ride to School)} \rightarrow \textit{(Ready by 8)}$$

And finally into:

$$S \rightarrow R$$

When we get to propositional logic, you'll learn how the arrow operator works. With that knowledge and the methods of logical analysis, we can prove that this statement implies the following:

If you aren't ready by 8, then you aren't getting a ride to school

Either you aren't getting a ride to school, or you'll be ready by 8.

And *doesn't* imply the following:

If you don't get a ride to school, that means you weren't ready by 8.

Why study logic? Well, there are many answers to that question, but here are a few I like the best:

1. Let's be honest, you'll never build a truth table again after leaving this class, so why waste our time doing it here? Good Question!

Because once we "look under the hood" of arguments by exploring their structure through logical analysis, we gain a deeper appreciation for what makes arguments tick and how arguments demonstrate their conclusions. This is just like understanding how a car goes. It's fine to go without much knowledge of what happens under the hood of a car until something goes wrong and you need to fix the car yourself. You don't want to end up stranded!

2. We also begin to internalize the distinctions between premises and conclusions, truth and validity, validity and soundness, consistency and logical equivalence, and others. Over time working with these concepts, thinking in these terms becomes natural. Furthermore, we begin to internalize certain simple and common logical relations and logical forms. These become second nature and we naturally and habitually think in more precise ways.

3. Finally, we get a window into one of the key tools that philosophers use to understand and critique arguments and positions. We also open the door to more complex logical analysis, which is important to mathematicians, linguists, computer scientists and programmers, artificial intelligence researchers, and philosophers alike.

5.1: Core Concepts

Truth Preservation

An important concept in logic is Truth Preservation.

What does it mean? An inference is **Truth Preservative** if and only if its premises being true *guarantees* that its conclusion is true. Sound familiar? That's because this is what **Validity** means. Truth Preservation means moving from premises to the conclusion (or from the antecedent to the consequent in an implication), we can't "lose truth." If the premises are true, then the conclusion is true and we haven't "lost the truth" of the premises. The inference *preserves* truth.

This is important because a valid deductive inference will always be *truth preservative*.

For example:

All humans are mortal

Socrates was a human

Therefore Socrates was mortal

...is truth preservative, because if it's true that All humans are mortal (true by definition, I'd think) and if it's true that Socrates was human (a matter of historical fact), then "Socrates was Mortal" cannot possibly be false. The inference is truth preservative, so we can't "lose" the truth in the premises in our inference to the conclusion (we can't end up with a false conclusion).

It's as if you begin a journey in premise base camp. You start loaded up with truth supplied by the premises. If it's a good path from premises to conclusion, you still have truth when you get to the conclusion (the conclusion is true). More than that: you are *guaranteed* to still have the truth you start with. If it's a bad path, then you may lose that truth you started with and end up with a false conclusion. If you happen to arrive at a true conclusion, it's not guaranteed—it's a mere accident. Truth preservative argument structures guarantee that you will have a true conclusion if you have true premises.

Deductive vs. Inductive

Remember also that we need to be able to distinguish between **deductive** and **inductive** arguments.

- Deductive arguments:
  - Formally precise arguments, **necessarily true**
  - Mathematical, logical, from definition,
- Inductive arguments:
  - Informal support, **probably true**.
  - Prediction, analogy, generalization, authority, causal inferences

Review section 1.2 in this textbook for a refresher on the difference between Deductive and Inductive arguments or inferences.

We will focus exclusively on deductive arguments for the next 2 chapters.

Form vs Content

Logic is often called **Formal Logic** because it is the study of how arguments work at the level of their *form or structure*. Many arguments can share one structure. Here's an example:

No Pineapples are Trees

So, No Trees are Pineapples

Now let's take the content away and just look at the structure:

No [Math Processing Error] are [Math Processing Error]

So, No [Math Processing Error] are [Math Processing Error]

We can see, since we've removed the terms (the content) and replaced them with symbols, what the structure of the argument is. Variables stand in for content in that they represent any possible term or proposition or the like that could be plugged into the space they occupy.

So, we can now produce a new argument with the *same structure*, by plugging in new terms.

No Friends are Enemies

So, No Enemies are Pineapples

Oooops! Missed a spot.

No Friends are Enemies

So, No Enemies are Friends

Can you see how this argument has *new content* while having the *same structure* as our inference about pineapples and trees from above? Logic is the study of argument structures and so when translating into a logical language, we'll sometimes just replace the content with variables because we no longer care if the argument is about Pineapples or Friends or Quantum Fields!

Great, let's move on to Categorical Logic:

CHAPTER OVERVIEW

6: Categorical Logic

This "chapter" is more or less a set of guides that Andrew Lavin created. It may not work as a standalone introduction to Aristotelian logic and so other resources may need to be used in conjunction. Treat this more as a supplement on Aristotelian Logic than a full introduction.

6.1: The Basics

Categorical logic was devised by Aristotle (384-322 BCE) and developed throughout Western history all the way up until the 19th Century. This was the dominant system really up until the early 20th Century.

Categorical logic concerns the relations between **Categorical Propositions**, which are propositions that relate categories of things.

For example:

> ✔ Example [Math Processing Error]
>
> *All cats have fur*

Is a false categorical proposition, but is a categorical proposition nonetheless.

> ✔ Example [Math Processing Error]
>
> *Some cats are things without fur*

Is a true categorical proposition, since Sphynx Cats don't have fur (or at least have no fur coat).

Each of these relates the category *cats* with some other category like *things with fur* or *things without fur*. Some of these relations obtain in the real world (are true or real relations) and some of them do not obtain (they are false or fake or illusory relations). *Cats*, it turns out, don't all have fur and so the relation of "all ____ are ____" doesn't hold between *cats* and *things with fur*.

Standard Form

In Categorical Logic, we're trying to get propositions into **Standard Form** so we can analyze all categorical propositions using the same basic set of tools of analysis. Standard form is a tool we use. When I express propositions using standard form, I enable myself to understand that proposition and its relation to other propositions more clearly.

Standard form consists of four elements:

1) **Quantifiers**: "All," "No," or "Some"

2) **Subject Term**: Noun or Noun Phrase

3) **Copula**: "are" or "are not"

4) **Predicate Term**: Noun or Noun Phrase

So, for example, any given categorical proposition will look like one of the following:

> ✔ Example [Math Processing Error]
>
> **A:** *All* ***birds*** *are* things that have wings
>
> **E:** *No* ***chickens*** *are* things that can swim
>
> **I:** *Some* ***bracelets*** *are* very expensive things
>
> **O:** *Some* ***people named 'Andrew'*** *are not* dogs

The Categorical Propositions relate classes in different ways:

A) **Total inclusion**: the subject category (the first) is totally included in the predicate category (the second). Meaning: all members of the subject category are also members of the predicate category.

E) **Total exclusion:** *the subject category and the predicate category don't share any members in common.*

I) **Partial inclusion**: *at least one* member of the subject category is included in the predicate category.

O) **Partial exclusion**: *at least one* member of the subject category is *not* in the predicate category.

Euler Diagrams:

An Euler diagram is a diagram that uses bubbles to represent categories and their relations to one another and to individuals. There's no standard "empty" diagram, so the way we draw the bubbles is how we represent these relations. In contrast, we'll see that Venn diagrams have a blank diagram that is always the same and then we use shading and X's to represent the relationships between the categories.

Figure *[Math Processing Error]*: **A Proposition:** All cats are mammals

Figure *[Math Processing Error]*: **E Proposition:** No Frogs are Fish

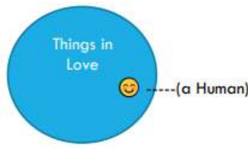


Figure *[Math Processing Error]*: **I Proposition:** Some humans are things that are in love

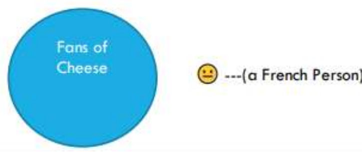


Figure *[Math Processing Error]*: **O Proposition:** Some French people are not fans of cheese

Quality and Quantity

Categorical Propositions have quality and quantity.

Quality: are they affirmative or negative? Do they posit inclusion or deny inclusion?

Quantity: are they about *all* things in the subject category? Are they "total" inclusion or exclusion relations? Or are they about *one or more* things in the subject category? Are they "partial" inclusion or exclusion.

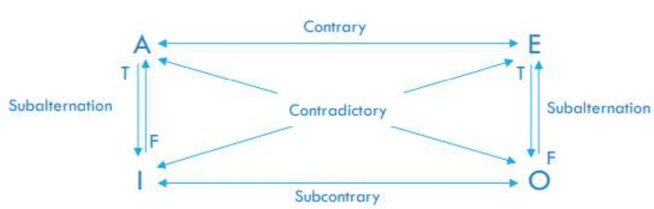
By answering these two questions (what is the quantity? and what is the quality?), we find out what sort of proposition we're talking about:

Table *[Math Processing Error]*

	(Quantity)	Affirmative	Negative
(Quality)			
Universal		**A: All Humans are Mortal**	**E: No Humans are Mortal**
Particular		**I: Some Human is Mortal**	**O: Some Human is not Mortal**

This table looks just like what is called the **Square of Opposition.**

Square of Opposition



The Square of Opposition tells us about the different relations between propositions *with the same subjects and predicates* translated into different forms.

A and E, as we can see, are **Contraries**, meaning they can't both be true.

Example [Math Processing Error]

FOR EXAMPLE, IT CAN'T BE THE CASE THAT ALL TEA CUPS ARE FRAGILE AND ALSO BE THE CASE THAT NO TEA CUPS ARE FRAGILE. IT COULD BE THAT SOME ARE FRAGILE AND SOME ARE NOT, SO THEY COULD BOTH BE FALSE. BUT THEY CAN'T BOTH BE TRUE.

A and O are **Contradictories**, meaning that they can neither be both false, nor can they both be true. One must be true and the other must be false. This is the purest opposition: they mean the exact opposite thing from one another.

E and I are also Contradictories.

Example [Math Processing Error]

FOR EXAMPLE, IT CAN'T BE THE CASE THAT ALL HUMANS ARE MORTAL, IF ONE HUMAN IS IMMORTAL (I.E. NOT MORTAL). SIMILARLY, IT CAN'T BE THE CASE THAT NO CHICKENS CAN SWIM, IF ONE CHICKEN CAN SWIM.

EITHER AT LEAST ONE HUMAN IS NOT MORTAL OR ALL HUMANS ARE MORTAL—THEY CANNOT BOTH BE TRUE AND THEY CANNOT BOTH BE FALSE.

EITHER AT LEAST ONE CHICKEN CAN SWIM OR NO CHICKENS CAN SWIM—THEY CANNOT BOTH BE TRUE AND THEY CANNOT BOTH BE FALSE.

WHY? WELL, *WHAT IT MEANS* FOR IT TO BE FALSE THAT ALL HUMANS ARE MORTAL IS SIMPLY FOR THERE TO BE SOME HUMAN THAT IS NOT MORTAL. SIMILARLY, *WHAT IT MEANS* FOR IT TO BE FALSE THAT NO CHICKENS CAN SWIM IS SIMPLY FOR THERE TO BE A CHICKEN THAT CAN SWIM.

I and O, as we can see, are **Subcontraries**, meaning they can't both be false.

Example [Math Processing Error]

FOR EXAMPLE, IT CAN'T BE FALSE THAT SOME PUPPIES ARE RAMBUNCTIOUS AND ALSO FALSE SOME PUPPIES ARE NOT RAMBUNCTIOUS.

TO SEE THIS, THINK ABOUT THE CONTRADICTORIES OF EACH. IF IT'S FALSE THAT I AND IT'S FALSE THAT O, THEN THE CONTRADICTORIES MUST BE TRUE. IN THIS CASE, THAT MEANS THAT ALL PUPPIES ARE RAMBUNCTIOUS AND ALL PUPPIES ARE NOT RAMBUNCTIOUS. THAT IS CLEARLY NONSENSE!

Subalternation is strange, but we'll go over it quite quickly. If the particular proposition is false, then the corresponding universal proposition is false. If the Universal proposition is true, then it follows that the corresponding particular proposition is true.

Example [Math Processing Error]

FOR EXAMPLE, IF ALL LEGO PIECES ARE RED, THEN IT FOLLOWS THAT THERE IS AT LEAST ONE RED LEGO PIECE (OBVIOUSLY, RIGHT?)

IF THERE ARE NO NEON GREEN LEGO PIECES, THEN IT FOLLOWS THAT AT LEAST ONE LEGO PIECE IS NOT NEON GREEN (LESS OBVIOUS, BUT STILL PRETTY CLEAR).

GOING THE OTHER WAY...

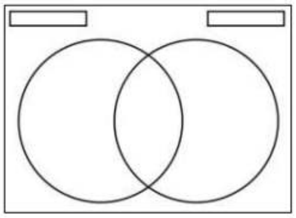
IF IT'S FALSE THAT SOME DAISY IS PURPLE, THEN IT *MUST* BE FALSE THAT ALL DAISIES ARE PURPLE.

IF IT'S FALSE THAT AT LEAST ONE ROBOT IS NOT A CONSCIOUS BEING, THEN IT *MUST* BE FALSE THAT NO ROBOTS ARE CONSCIOUS BEINGS (AFTER ALL, AT LEAST ONE ROBOT IS A CONSCIOUS BEING!)

6.2: Venn Diagrams

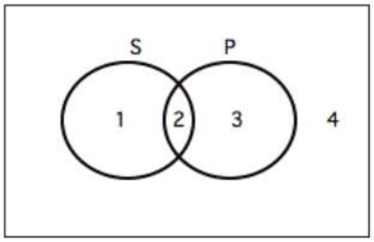
A Venn Diagram in some ways improves on an Euler diagram by having a blank form that we can shade and X to create each of our propositions. We'll see later how this allows us to do our analysis of inferences.

Blank form:

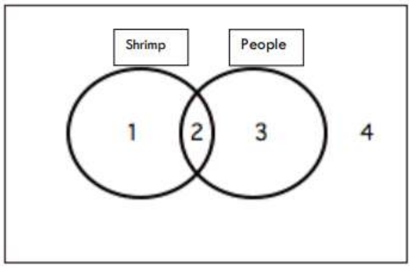


We have two overlapping circles, labels for each circles, space outside of both circles, and room for putting X's and shading in all of these regions.

Let's put number labels on each of these regions:



Let's make this more concrete so we can discuss each region.



1) Is the region of Shrimp that are not People.

2) Is the region of things that are both Shrimp and People (you won't find any).

3) Is the region of things that are People and not Shrimp.

4) Is the region of non-shrimp and non-people (Tubas, Cars, Bananas, etc.)

Okay, so how do we diagram each of our standard form categorical propositions?

The first rule to note is that all universal propositions are diagrammed using **Shading** and all particular propositions are diagrammed using **X's**.

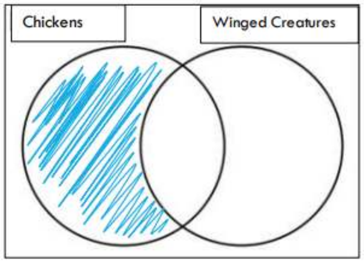


Figure *[Math Processing Error]*: A: All Chicken are Winged Creatures

We shade in region 1 because there are no Chickens that are not Winged (at least that's what our proposition claims).

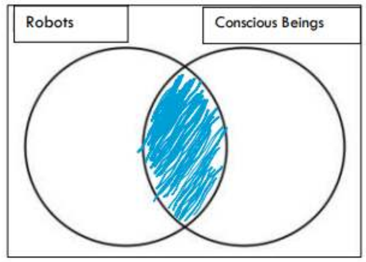


Figure [Math Processing Error]: E: No Robots are Conscious Beings

We shade in region 2 because the proposition claims that there are no Robots that are also Conscious Beings.

NOW WE SWITCH TO X'S INSTEAD OF SHADING

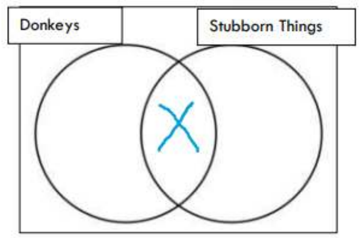


Figure [Math Processing Error]: I: Some Donkeys are Stubborn Things

We put one X in region 2 because the proposition claims that there is at least one Donkey that is also a Stubborn Thing.

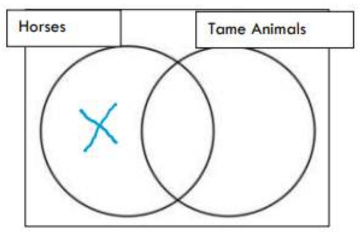


Figure [Math Processing Error]: O: Some Horses are not Tame Animals

We put one X in region 1 because the proposition claims that there is at least one Horse that is not a Tame Animal.

Using Venn Diagrams for Inferences

Any immediate inference will be apparent in a Venn diagram.

What's an immediate inference, you ask?

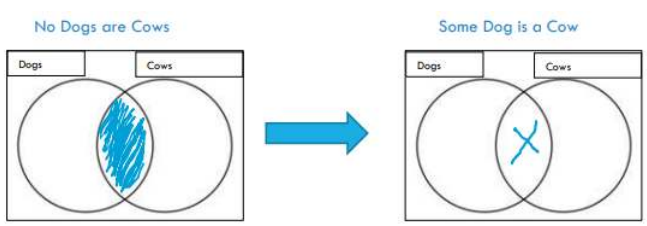
It's an inference from one proposition directly (immediately) to another inference.

Let's look at the following inference:

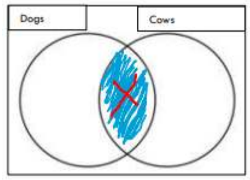
No dogs are cows, so it follows that

it is false that there is some dog that is a cow.

Compare our two diagrams:



Do you see how, if we overlapped them, they wouldn't fit together?



This is how we know that if the first proposition is true, the second one must be false.

These are **Contradictories**, so of course if one is true, then the other is false.

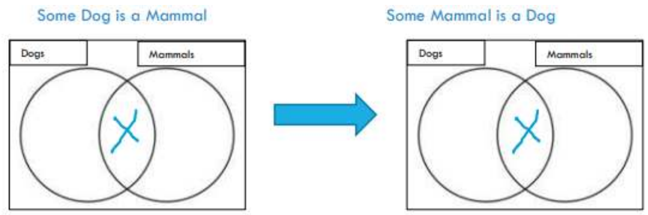
This inference, therefore, doesn't work. It's an **Invalid** inference.

Let's try another:

Some dog is a mammal,

so it follows that some mammal is a dog

Notice how the two diagrams are identical:



If they're claiming the exact same thing, then obviously one implies or entails the other. What it means for each proposition to be true is for the same state of the world to obtain (to be the case).

Therefore, this inference is **Valid**.

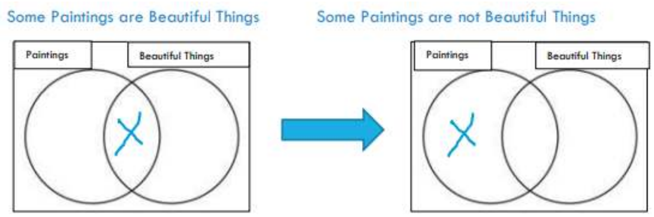
This type of inference (switching the subject and predicate) is called **Conversion** and it only works for E and I Propositions. Try to do it for an A or O proposition and you'll find that it doesn't work.

And another type of inference that often arises:

Some paintings are beautiful,

so it follows that some paintings are not beautiful.

Let's look at the two diagrams.



Notice how the information contained in the second diagram *is not contained in the first diagram*. Immediate inferences only work (and in fact all deductive inferences only work) if the information contained in the conclusion is already contained in the premises.

> *If in the move from premises to conclusion, we have to change our diagram (add new information), then the inference is* ***invalid***
>
> *(at minimum there is a hidden assumption,*
>
> *at worst the inference simply doesn't work).*

So, to recap, there are **three kinds** of inference here:

1. Invalid inferences because the diagrams are incompatible.

2. Valid inferences because the diagrams are the same (or sometimes because the conclusion is represented in the diagram in some way).

3. Invalid inference because the conclusion introduces new information (changes the diagram).

Conversion, Obversion, Contraposition

Let's talk briefly about three kinds of immediate inferences one can draw in Categorical Logic. These are in addition to the inferences we can draw using the square of opposition.

Conversion

Conversion is the switching of the subject and predicate.

So, "No Apples are Bananas" converts to "No Bananas are Apples."

And, "Some Frisbees are round things" converts to "Some round things are frisbees."

This only works for E and I propositions. We can remember this by looking at the middle vowels of the word Conv**E**rs**I**on.

Obversion

Obversion is when you **change the quality** of the proposition from negative to affirmative or affirmative to negative, and then you **replace the predicate with its complement.**

So, "No Apples are Bananas" obverts to "All Apples are non-Bananas."

And, "Some Frisbees are round things" obverts to "Some Frisbees are not non-round things."

Obversion works for every form of categorical proposition.

> **Definition: Complements**
>
> What's a complement, you ask?
>
> A complement of class/kind/category X is whatever isn't a member of X.
>
> The complement of **birds** is the set of all **non-birds**.
>
> The complement of **people identical to Lakshmi** is the set of all **things not identical to Lakshmi.**
>
> The complement of **animals that are going to get eaten** is the set of all **things that aren't animals that are going to get eaten**.
>
> We form complements by introducing a negation (non, not, n't, etc.) into the term. Usually, you can just put a "non-" in front of the term, but, as you can see, that isn't always easy.
>
> The important thing to remember is that your description should cover *everything* that isn't in the group (all non-existent things, all imaginary things, all real things, everything). So don't make the complement of "people identical to Sven" by writing "People not identical to Sven."

Contraposition

Obversion is when you **switch the subject and predicate** of the proposition, and then you **replace each with its complement.**

So, "All Apples are Bananas" obverts to "All non-Bananas are non-Apples."

And, "Some Frisbees are round things" obverts to "Some non-round things are non-Frisbees."

This only works for E and I propositions. We can remember this by looking at the middle vowels of the word Contr**A**p**O**sition.

6.3: Categorical Syllogisms

A **Syllogism** is a two-premise inference. A **Categorical Syllogism** is a two-premise inference where the premises and the conclusion are categorical propositions.

For example:

Example [Math Processing Error]

All Smart Animals can Use Computers

All Monkeys are Smart Animals

Therefore all Monkeys can use Computers.

I've deviated from standard form a bit (can you see how?), but we can see the basic anatomy here. The first premise is false, btw.

We have three "terms" here. Subjects and Predicates of each proposition are different kinds of "terms" or categories that we're dealing with.

1) **Major Term**: The term that appears in the first premise and the conclusion as its predicate.

2) **Middle Term**: The term that doesn't appear in the conclusion (only in both premises)

3) **Minor Term**: The term that appears in the second premise and the conclusion as its subject.

So in our example here, The Major Term is *Things that can use computers*, the Minor Term is *Monkeys*, and the Middle Term is *Smart Animals.*

Finally, the premises have names as well:

1) **The Major Premise** is the first premise (must have the major term)

2) **The Minor Premise** is the second premise (must have the minor term)

Mood and Figure

Syllogisms also have Mood and Figure.

Mood is simply a list of the categorical proposition forms used in the argument. Our argument about monkeys above is an AAA syllogism because it uses three A propositions. You can also have EIO or EAE or OOO syllogisms. The order is always Major Premise, Minor Premise, Conclusion. This is pretty straightforward.

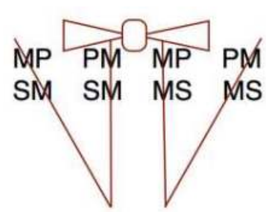
Figure is a description of the "shape" of the proposition based on where the middle terms fall.

There are four figures:

Figure 1	Figure 2	Figure 3	Figure 4
[Math Processing Error]	*[Math Processing Error]*	*[Math Processing Error]*	*[Math Processing Error]*

Above, we have the figures with their labels. Note that "P" means "major term" (or predicate of the conclusion) and "S" means "minor term" (or subject of the conclusion).

Below, you'll see them laid out graphically, and you can see how they make the shape of a dress shirt's collars. This is a mnemonic device for remembering the shape of the figures.



Once we know the Mood and Figure of a syllogism, we can a) reconstruct the structure in standard form without any more information and b) identify whether or not it's a valid syllogism structure. How? Well by memorizing the list of valid structures.

Three-Term Diagrams

Okay, now that we've got some of the terminology out of the way, we can start to see how syllogisms work.

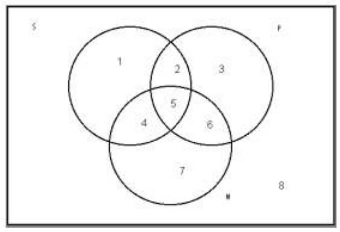
No Contradictions are True Statements

All True Statements are Trustworthy Statements

Therefore No Contradictions are Trustworthy Statements

Syllogisms post relations between categories that, taken together, *already contain* another relation between a category. The best way to see this is by going straight to Venn Diagrams.

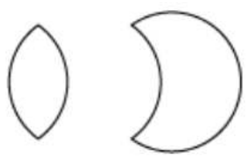
In order to diagram *three terms*, we'll need to beef up our Venn Diagram a bit:



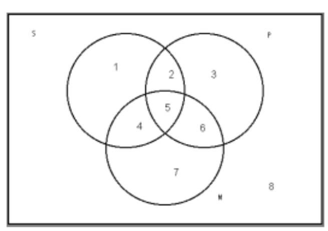
The names of the regions aren't super important. You'll find different numbers or letters being used in different texts and videos. What is important is that we see that we now have **three circles** and more regions that can be filled in or X'd.

Rest assured, though, we're still only ever working with two circles at a time, just like we did with immediate inferences and single propositions.

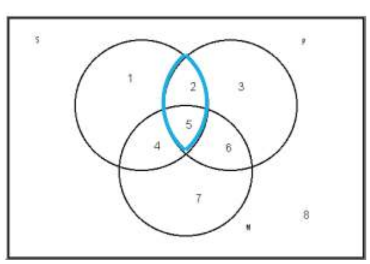
This means that you'll only ever be focusing on a region that looks like these:



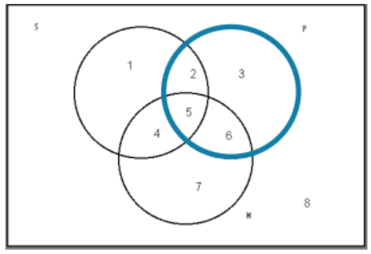
So even if your diagram looks like this:



You'll still focus only on this:



Or this:



At any given time.

Analyzing Syllogisms

Okay, how do go about analyzing syllogisms?

Let’s recall our previous example:

Example *[Math Processing Error]*

No Contradictions are True Statements

All True Statements are Trustworthy Statements

Therefore No Contradictions are Trustworthy Statements

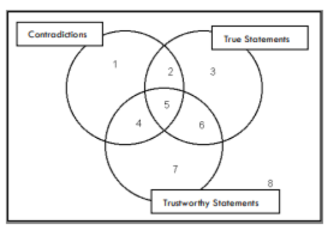
We have three terms in this syllogism:

1) Trustworthy Statements (Major Term)

2) True Statements (Middle Term)

3) Contradictions (Minor Term)

We assign each a circle (just like we have been doing):



Then we take each statement and diagram it normally:

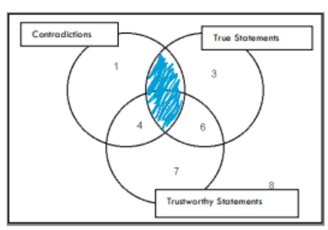
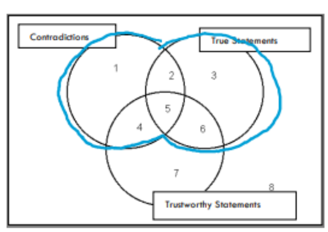


Figure *[Math Processing Error]*: No Contradictions are True Statements

Note that here, we’re only focusing on these two circles and ignoring the other:



Moving on...

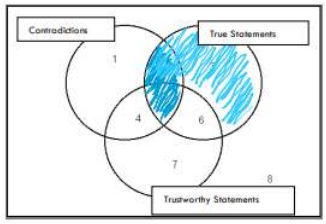


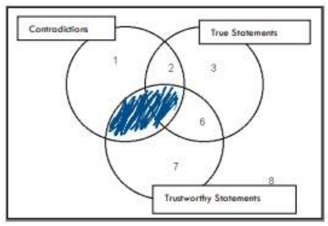
Figure *[Math Processing Error]*: All True Statements are Trustworthy Statements

This time, we've focused on a different pair of circles and ignored Contradictions instead.

Now we look at the conclusion and try to discern whether the information is already present in the diagram:

Therefore, No Contradictions are Trustworthy Statements

Here's the diagram for this statement:



But this isn't contained in twice-shaded the diagram above because region 4 is not totally filled in. For all we know, according to our twice-shaded diagram, there could be a Contradiction that is a Trustworthy Statement (there could be something in region 4).

Therefore, this syllogism is **Invalid**.

Here's another example from above:

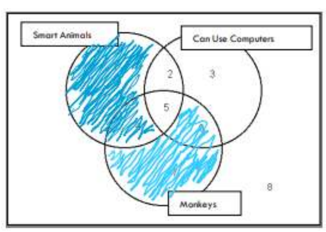
Example *[Math Processing Error]*

All Smart Animals can Use Computers

All Monkeys are Smart Animals

Therefore all Monkeys can use Computers.

If we diagram this syllogism following the steps above, we get the following diagram:



Then we look at our "target diagram" or the diagram of the conclusion:

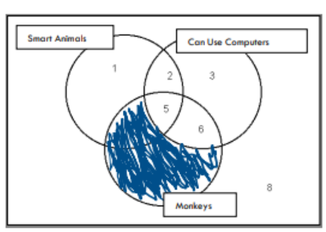
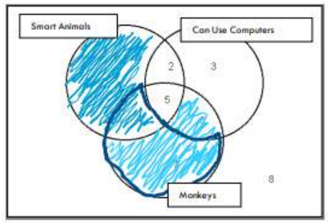


Figure *[Math Processing Error]*: Therefore all Monkeys can use Computers.

And we find that *it is already represented in our premise diagram*:



Notice how the dark blue outline is totally filled in in the premise diagram, so in overlaying our conclusion diagram, *we won't be changing the diagram at all.*

Therefore, this syllogism is **Valid**.

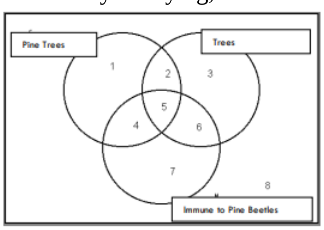
Let's try a particular syllogism:

No Pine Trees are Immune to Pine Beetles

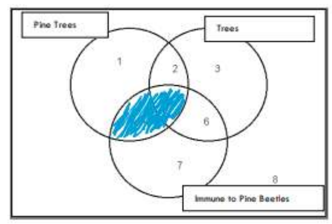
Some Tree is a Pine Tree

So, sadly, Some Tree is not Immune to Pine Beetles

Again, I've deviated from standard form, but so long as we understand what they're saying, it isn't too important.

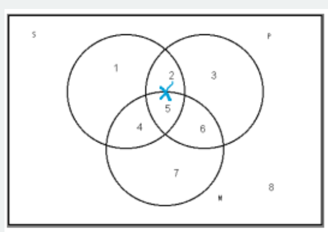


Here we are after the first premise is diagrammed:



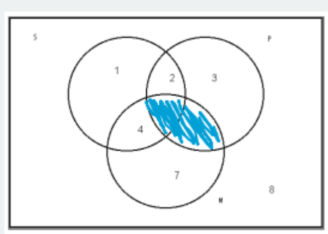
Now, normally when we put an X in a diagram, we'll do the following:

Say the proposition is Some S is a P



Note how I've put the X **on the line**, signifying that we don't know whether it goes in region 2 or 5.

Now Suppose we already have the proposition No M is a P in place:

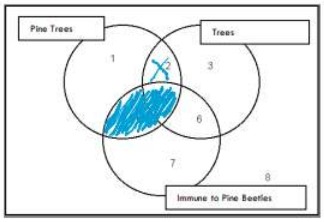


This time, when we go to diagram Some S is a P, we'll be forced to put it in region 2 since we know it can't be in region 5.

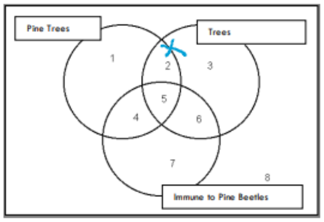
In general, therefore, it's best to **diagram the universal statements first.**

Okay, back to Pine Trees:

Piney is a Pine Tree, so we put an X:



Then we ask whether the conclusion diagram is already represented.



Note how our X is straddling the line between 2 and 3 because any tree in 2 or 3 is not Immune to Pine Beetles, which is what we're trying to express.

Is that information contained in the previous diagram? Yes, there is an X in 2, which is more specific than "There is an X in 2 or 3". We know, therefore, that there is a tree that isn't immune to pine beetles and that tree is one of the pine trees. So, this argument is **Valid.**

6.4: Proving the validity of immediate inferences

Let's take a second to look at the different immediate inferences using Venn diagrams. Why do they work? Why don't the invalid inferences work?

Square of Opposition

Contradictories

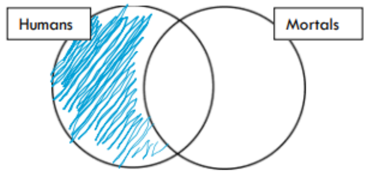


Figure *[Math Processing Error]*: A: All humans are mortal things

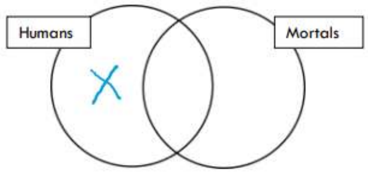
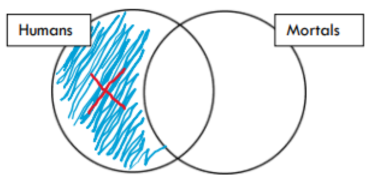


Figure *[Math Processing Error]*: O: Some humans are not mortal things

Taken together:



There can't be an X in a shaded area. We've "filled it with concrete" so there's nothing there. What does this mean? These two propositions are claiming the opposite thing from one another. A is claiming that there are No things in region 1 (the left moon-shaped region), whereas O is claiming that there is at least one thing in region 1.

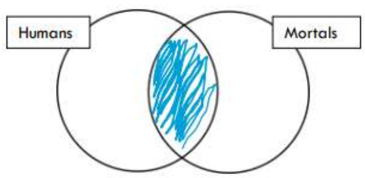


Figure *[Math Processing Error]*: E: No humans are mortal things

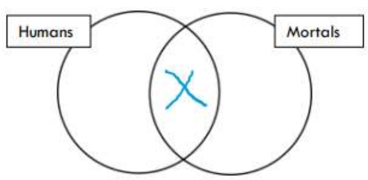
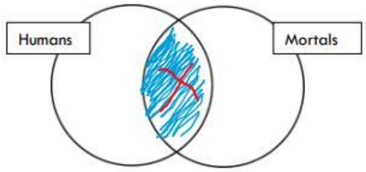


Figure *[Math Processing Error]*: I: Some humans are mortal things

Taken together:



Again, E is claiming that there is nothing in region 2 (the ellipse or football shape in the center) and I is claiming that there's at least one thing in region 2.

Contraries

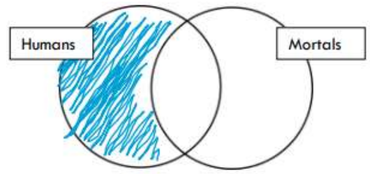


Figure [Math Processing Error]: A: All Humans are Mortal

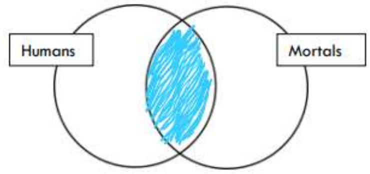
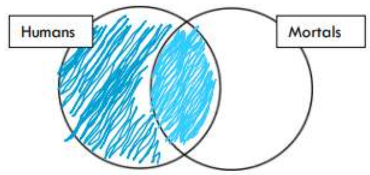


Figure [Math Processing Error]: E: No Humans are Mortal

Taken together:



Remember that these are only contrary on Aristotle's assumption that the subject class has at least one member. Notice how in this diagram, the subject class (humans) must be empty. It is claiming that there are no humans. Since Aristotle assumes that the subject class is populated by at least one thing, it follows that A and E propositions cannot both be true (they are contraries).

I won't do subcontraries or subalternation because they're more complicated and hopefully you get the idea by now.

Other Immediate Inferences

Conversion

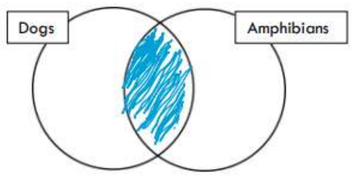


Figure [Math Processing Error]: E: No Dog is an Amphibian

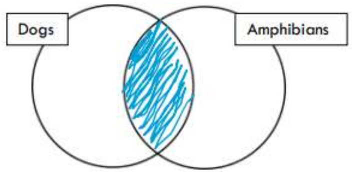


Figure [Math Processing Error]: E*: No Amphibian is a Dog

See? They're the same diagram! So the immediate inference is valid: we're claiming the same thing in each proposition.

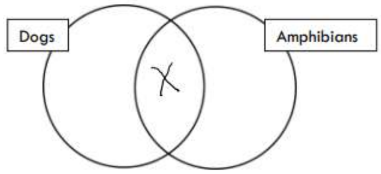


Figure [Math Processing Error]: I: Some Dog is an Amphibian

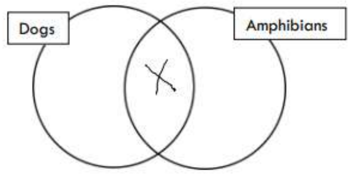


Figure *[Math Processing Error]*: I*: Some Amphibian is a Dog

Again, they look the same since they're both claiming that there's at least one thing in region 2.

But Conversion doesn't work with A's and O's:

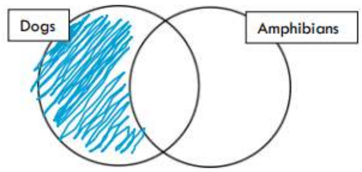


Figure *[Math Processing Error]*: A: All Dogs are Amphibians

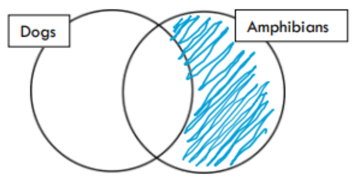


Figure *[Math Processing Error]*: A*: All Amphibians are Dogs

Notice how they are claiming almost opposite things. At any rate, A* is certainly not contained in the diagram for A. (That's what validity looks like, the conclusion is already in the diagram for the premises).

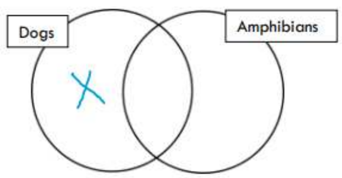


Figure *[Math Processing Error]*: O: Some Dogs are not Amphibians

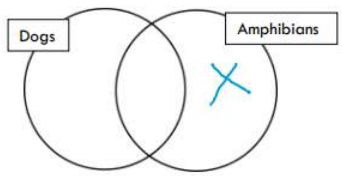


Figure *[Math Processing Error]*: O*: Some Amphibians are not Dogs

Again, they are claiming basically the opposite thing from one another and as a result, O* is not already represented in the diagram for O.

Obversion

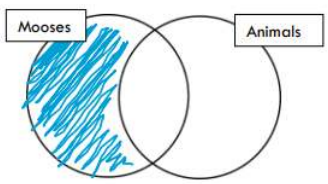


Figure *[Math Processing Error]*: A: All Mooses (Meese?) are Animals

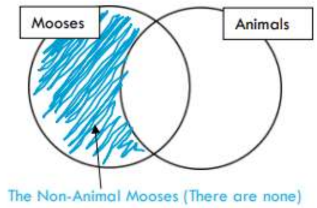


Figure *[Math Processing Error]*: A*: No Mooses are Non-Animal

There are no mooses in the non-animal group, which is the left moon-shaped region, or region 1. So we shade it in.

All of the other Obversions end up with the same venn diagram too. Check it. Then give Contraposition a shot!

6.5: Key Terms

- **Categorical Propositions**
- **Truth Preservation**
- **Form and Content**
- **Standard Form**
- **Conversion**
- **Obversion**
- **Contraposition**
- **Complement**
- **Mood**
- **Figure**

CHAPTER OVERVIEW

7: Propositional Logic

7.1: Logical Entailment

If I was to tell you that either I was going to be promoted or get a big bonus because I just landed a huge account for the company I work for, then you would think a few things:

- If he doesn't get promoted, then at least he'll get compensated with a bonus
- If he doesn't get a bonus, then at least he'll get compensated with a promotion
- His company won't both *not* promote him *and not* give him a bonus.

All of these claims follow from the original claim. They "follow" in the sense that if the original claim is in fact true, then this conclusion *must* be true. There's some sense in which they "mean the same thing": they describe the same world or claim the same thing to be true about the world. A lot of logic consequence is similar: it's a relationship of "following" or "entailment" between statements which mean essentially the same thing.

Other logical consequences are cases where one statement entails another statement (if the first is true, the second *must* be true), but not because they essentially mean the same thing. Instead, because the first statement is making a "stronger" claim than the other. Here's an example:

Example [Math Processing Error]

Franklin is a student at Butte College.

Entails the following claim:

Either Franklin is a student at Butte College or Franklin is a student at De Anza College, I'm not sure which.

This first statement entails this second statement because the second one is an "either...or..." statement and is therefore much weaker. If I told you "Either Franklin is a student at Butte College or Franklin is a student at De Anza College", then you'd know that *one of two* worlds must be the real world: either a world in which Franklin is a student at Butte or a world in which Franklin is a student at De Anza. Alternatively, if I told you "Franklin is a student at Butte College," then you'd know *which* of those two worlds was the real world: the one in which Franklin was a student at Butte College. It is therefore a stronger or more determinate claim.

Think in terms of betting: Which would you bet on: that *either* Russia or China would win the most Gold medals at the next Olympics, or that *Russia* would win the most? If you're wise to the logic of the situation, you'd know to bet on the *disjunction* or "or" statement since you'd have a higher change of winning. This is what we mean by "stronger claim" and "weaker claim": a stronger claim is a claim that is "harder to make true" in the sense that the world has to be one particular way to conform to that claim. A weaker claim, conversely, is relatively easier to make true in that the world could be a couple of different ways for the claim to be true.

Logic is the science of logical consequences: these relationships between statements such that if one is true, then the other *must* be true.

But logic consequence isn't like just any consequence: think about the following example:

Example [Math Processing Error]

If the Russians invaded the United States,

then the US would have to defend itself.

The second statement: that the US would have to defend itself, seems to *follow* in some sense from the first statement. Nevertheless, this is not a *logical* consequence: it's not an entailment that holds (or is valid) in virtue of the *structure* or *form* of the two statements involved. It's an inferential relationship that these propositions have in virtue of *the way the world works* or *the way war and international relations works*. That, also, isn't strict entailment. Logical entailment is deductive entailment in that if the premise(s) is true, the conclusion *must* be true. It couldn't possibly be false given the truth of the premise(s). For instance, if it is in fact true that:

If one gets properly poisoned, then one will get sick.

That is, if one can't possibly be poisoned without getting sick, then it follows *by logical entailment* that:

If one doesn't get sick, then one wasn't properly poisoned.

And it also follows that:

Either one didn't get properly poisoned, or one will get sick.

These are logical entailments in that these inferences hold *in virtue of form*. Any statement with the logical form:

If A, then B

Will entail corresponding statements with the corresponding logical forms of:

If not B, then not A

And

Either not A or B.

These are relationships that hold between *any statements at all* that have these logical structures. If the first one is true, then the other ones are necessarily true. Logical entailment is entailment that holds because of the logical form of the statements involved.

7.2: Propositions and their Connectors

Wouldn't it be great to have a system that tracked *all* of these logical entailments? Maybe we could devise a way of taking English sentences (and sentences in other natural languages as well!) and stripping them down to just their logical structures so that we could explore the properties of just the logical structures of statements and arguments.

And since the Greek word for form or structure is "logos", let's call this system "Logic", yeah?

Okay so Logic is the study of the logical forms of statements, sets of statements, and arguments. Propositional Logic is just a particular sort of logic. What makes it special?

Aristotelian Logic was discussed in the previous chapter. There, we focused on "terms" or "categories" and the various ways they can be related to one another. One category might totally include another, partially include another, or totally or partially *exclude* another. Here are some examples:

- The category "mammal" totally includes the category "dog", meaning that every dog is included in the category mammal—or simply: every dog is a mammal.
- "Animal" partially includes "cats" in that some animals are cats, but not all of them are.
- "Gem" partially excludes "diamond" because not all gems are diamonds: there are some gems that are not diamonds.
- "Human" totally excludes "dinosaur" in that no humans are dinosaurs and no dinosaurs are humans.

These are relations between *categories*, hence the name "Categorical Logic." Propositional Logic, alternatively, focuses on the relations between different entities: propositions or statements. Sometimes it is called "Sentential Logic" instead because it is the logic of how *sentences/statements/propositions* relate to one another.

If we're going to represent the logical form of a statement in English, we need to decide first *what* we want to symbolize (represent using symbols) and then second *how* we are going to symbolize those things. That is, we must decide what aspects of ordinary language is part of the logical form or is *logically relevant*. Once we have made that decision, we can symbolize the logically relevant bits and then leave the rest out. We'd then be left with *only* the logically relevant aspects of the original statement.

Once we have done this symbolization, we can then study the logical form on its own apart from the particular content of the statement we're interested in.

We'll learn 5, yes *five* different logical symbols and that's all. It's a bit like learning a language with only five words: translation will be a bit tricky sometimes, but you don't have to spend much time developing your vocabulary.

As a sneak preview, the five logical symbols that we'll learn correspond roughly to the following English words or phrases:

And

Not

Or

If...then...

If and only if...will...

First, since the task here is to learn about how to symbolize an English sentence into propositional logic—how to extract the logical form from a sentence, that is—we should learn the most basic operation: symbolizing an atomic proposition.

Atomic propositions are simple propositions that *have no logical content*. Here are some examples:

Example [Math Processing Error]

I've been swimming every day this week.

Example [Math Processing Error]

There is a bird nesting in the attic.

Example [Math Processing Error]

The Order of the Phoenix is a secret order from the Harry Potter book series that sought to undermine the goals of Voldemort.

Atomic propositions are sometimes longer, sometimes shorter, but they always have one thing in common: they make one unified a claim about the world that might be true and might be false. There are no relations between propositions (relations like ‘and’, ‘or’, or ‘if...then...’) nor are there any negations or ‘not’s’—in short, there is no logical content in an atomic proposition.

To translate an atomic proposition is quite simple. Replace one whole atomic proposition with a single variable called a “sentence letter”. Here is how you would symbolize the previous examples:

$$S$$

$$B$$

$$O$$

I just chose what I took to be the **most significant word** from each proposition and then used the first letter of that word as the variable. So I chose “swimming”, “bird”, and “Order”. We do this so it’s easier to tell which variable symbolizes which English sentence.

That’s all there is to symbolizing a simple or atomic proposition.

“And” AKA Conjunction

Ever heard of “conjoined twins”? These are twins that are stuck together. That name comes from the word “conjunction” or “conjoint” (remember conjoint premises from Chapter 3?). A conjunction is a sticking together or, simply, an “and”. Imagine the following exchange:

I am the director of marketing

Nice to meet you, I’m from an investment firm. We’re very interested in your company.

In that case: full-disclosure, I also work for the investment firm that owns a majority share in this company.

Wait, you’re the director of marketing and you work for their investment firm?

This final reply is a *conjunction*: a joining together of two otherwise separate statements. It makes perfect sense what an “and” or conjunction does: it says “both A and B are true.” The statement as a whole turns out to be false if either of the conjoined propositions turns out to be false.

Here are a few examples of conjunctions:

Example [Math Processing Error]

I’ve always been a quiet person, but when I get on stage I like to be loud.

Example [Math Processing Error]

No one will ever love you and you’ll die alone!

Example [Math Processing Error]

Lincoln was shot in Ford’s Theatre, but Kennedy was shot in a Lincoln Continental made by Ford. You do the math.

The first statement is only true if it is true that A: This person has always been a quiet person, and B: When this person gets on stage they like to be loud. If either of those statements ends up being false, then the whole statement above is false.

The second statement is only going to be true if A: no one will ever love the person they’re talking to, and B: the person they’re talking to will die alone. If someone ends up loving them but they die alone anyways, the second statement above is still false since one of the *parts* or *conjuncts* is false.

> **Definition: Conjunct**
>
> One part of a conjunction—one of the conjoined statements—is called a **"conjunct"**

How do we symbolize a conjunction? First, you'll need to learn to recognize the indicator words that tell us that we're dealing with a conjunction. Some really common ones are:

And

But

However

Yet

Remember from grammar or English class the categories "coordinating conjunction" and "subordinating conjunction"? Well a logical conjunction is usually marked by an English word from those categories.

Once we've recognized that we've got a conjunction, then we remove that indicator word and replace it with a conjunction symbol. There are two common conjunction symbols and we're going to learn both so that if your instructor shows you a video or sends you to a website that uses a different one, you'll be aware that different people use different symbols. Here are the two conjunction symbols:

[Math Processing Error]	*[Math Processing Error]*

The one on the left is called a *dot* and the one on the right is called a *wedge*. There is **no difference in meaning between the two**. They mean the exact same thing. You'll have to train yourself to see them as equivalent. A lot of logic requires training yourself in these ways—it's like learning a new language!Okay, so now that we have our symbols, we can work through an example:

> **Example** *[Math Processing Error]*
>
> *I'm not much of a romantic, but I wanted to ask for your hand in marriage.*
>
> **Solution**
>
> There's a bit more going on here than a conjunction, as you'll see when we work through the next symbol, but for now we only know about conjunctions and so we'll only symbolize the conjunction.
>
> The first step is to identify the indicator word—the word that indicates that we're dealing with a logical conjunction. Can you find it?
>
> Yep! It's the "but". We simply replace that word with our symbol:
>
> *I'm not much of a romantic [Math Processing Error] I wanted to ask for your hand in marriage.*
>
> And then we take out the rest of the content and replace each simple proposition with a **Sentence Letter**—an upper-case letter that stands in as a variable for any simple proposition.
>
> Because the only thing we know how to translate at this point is the "but", we'd remove the rest and replace it with variable letters:
>
> N *[Math Processing Error]* M
>
> Why did I choose "N" and "M"? I chose the N for "not" as in "not much of a romantic" and the M for "marriage". You simply choose a significant word from the simple proposition and then use that as the variable letter. There is a bit more too it than that in that you have to watch out for a few things, but we'll discuss that after a bit more on the other symbols.
>
> The last thing to do is to add parentheses around the outside. Every dot (*[Math Processing Error]*) or wedge (*[Math Processing Error]*) gets its own set of parentheses. Again, there's a grammar to propositional logic, but we'll discuss that in due time.
>
> (N *[Math Processing Error]* M)
>
> There, we've done it! We've noted that this sentence is a conjunction between two simple propositions in terms of its logical form.

“Not” AKA Negation

Not so fast! Isn’t weird to use “N” to stand for “not romantic?”

Good catch, student, good catch. It feels a bit weird to translate “I’m not romantic” as a simple proposition when “I’m romantic” feels like the simple proposition that’s being *denied* here. It would be good if we could capture the relationship between “I’m romantic” and “I’m not romantic” using a logical symbol. Otherwise we just translate them as simple propositions and the result is just R and N. There’s no clear logical relationship between R and N.

Maybe we could introduce a logical symbol that stands for “not” so we can capture this relationship? Let’s use these two symbols:

~	*[Math Processing Error]*

The left one is called a *tilde* and it’s on the top left of your keyboard. The right one is called a *hook* and has to be found in the “insert-symbol” dialogue box in Microsoft Word. Again, different people use different symbols, but these are the two most common. They again mean the same thing as one another—there’s no difference between them for our purposes here.

Now, we’d take a simple proposition—a proposition without *any logical content* and symbolize it using a variable letter. So “I’m romantic” becomes “R”. Then if there’s a negation, a denial, or a “not” in the sentence, we’ll use a tilde or hook to symbolize the idea that there’s a logical negation in the ordinary language sentence:

*[Math Processing Error]*R

We might read this “not R”. It means something like “It’s not the case that the simple proposition R is true.” If we take one step back, we can see what the negation is doing a bit more clearly:

[Math Processing Error](I’m romantic).

We might read this “It’s not the case that I’m romantic.” I’m not romantic. The negation symbol marks for us that a simple positive proposition “I’m romantic” is being *denied* here.

Here are some examples of sentences with negations in them:

Example *[Math Processing Error]*

I’m not the person you think I am

*[**It’s not the case that** I am the person you think I am]*

Example *[Math Processing Error]*

The doctor isn’t in right now, may I take a message?

*[**It’s not the case that** the doctor is in right now]*

Example *[Math Processing Error]*

None of our kids have ever failed a class.

*[**Not** one of our kids has failed a class]*

Notice how in the last one the simple proposition “one of our kids has failed a class” sounds “negative” in that failing a class is a bad thing. Logical negation isn’t all types of “negativity.” Instead, logical negation is when a statement is being *denied*.

Also notice the second example here has an “isn’t” in it. The “not” is contracted. This makes it sometimes easy to miss a negation. Keep an eye out for “n’t” as in “mustn’t” “isn’t” “can’t” “shouldn’t” and of course “whomst’ven’t” ;).

Let us revisit our example from the previous section on conjunction:

I’m not much of a romantic [Math Processing Error] I wanted to ask for your hand in marriage.

Once we notice that there is a negation in the left conjunct (the left proposition), we can symbolize that negation *before* we eliminate the rest with a simple proposition variable:

[Math Processing Error] I'm a romantic [Math Processing Error] I wanted to ask for your hand in marriage.

Now that we're truly down to simple propositions with no logical content, we put our variables in. Again, we just choose what we think is the most significant letter in the sentence:

[Math Processing Error] R *[Math Processing Error]* M

Finally, remember that every conjunction (*[Math Processing Error]* or *[Math Processing Error]*) gets a set of parentheses and that those parentheses go *outside of each conjunct*:

(*[Math Processing Error]* R *[Math Processing Error]* M)

Remember that this symbolization would be equivalent since we have equivalent symbols:

(~ R *[Math Processing Error]* M)

Note also that you wouldn't want to symbolize it this way:

[Math Processing Error] (R *[Math Processing Error]* M)

Where the hook (*[Math Processing Error]*) is on the outside of the parentheses. This means something entirely different. Something like "I'm not both romantic and want to propose marriage." This is not what the original sentence said, so it'd be an inaccurate symbolization.

"Or" AKA Disjunction

We've now covered two logical symbols: *[Math Processing Error]* and *[Math Processing Error]* (and their equivalents *[Math Processing Error]* and ~). We can symbolize and/but/however (conjunction) and also not (negation). What happens if you see sentences like the following?

I'll either strike out this inning or get a runner into home base to score the winning run.

Look, we know that the Republican or the Democrat will win the race, so there's no point in voting for a third party.

The world will end unless we do something about global warming.

In each of these sentences, there is a *disjunction* between two possibilities presented. The idea seems to be that one of the options will happen no matter what. We will do something about global warming *OR the world will end. A Republican will win OR a Democrat will win. I'll strike out OR I'll score the winning run. The following symbol symbolizes a logical disjunction:*

[Math Processing Error]

One theory is that this symbol started because in Latin "vel" means "or" and so the "v" at the beginning became the symbol for "or" in propositional logic. Sounds plausible enough. This symbol is sometimes called a "vee" or simply a disjunction symbol. We won't learn an equivalent for it.

Disjunctions are similarly simple in that there are really only two words that it translates:

Or

Unless

Often the "or" is in a phrase using "either...or...", but that's basically it: don't see an 'or' or 'unless'? Then it's almost certainly not a disjunction.

Remember that a disjunction means that "at least *one* of the disjuncts is true".

> **Definition: Disjunct**
>
> One of the things being disjoined by a disjunction—one side of the disjunction—is called a **"disjunct"**

Translation for disjunction works almost the exact same way as conjunction. Identify the two propositions being connection via disjunction, then put the disjunction "*[Math Processing Error]*" between them and enclose in parentheses all the way outside each disjunct:

The world will end unless we do something about global warming.

(The world will end) *[Math Processing Error]* (*we do something about global warming)*

E *[Math Processing Error]* D

(E *[Math Processing Error]* D)

All done! Just remember that every "*[Math Processing Error]*" gets its own set of parentheses.

"If...then..." AKA Implication AKA Hypothetical AKA Conditional

If you've followed along well so far, then you'll do just fine in the propositional logic unit.

Notice how this sentence has an "if...then..." grammar. Since if...then...has a very clear logical definition, we can use a symbol to capture this structure and include it in our symbolization of the logical structure of this proposition. Put another way: we can capture the "if...then..." relationship using a logical symbol. Here are the symbols we'll use:

[Math Processing Error]	*[Math Processing Error]*

The left one is called an *arrow* (for obvious reasons) and the right one is called a *horseshoe* (for still fairly obvious reasons). The right one comes from a specialized logic used primarily in mathematics called "set theory". As a result, many people who teach intro to logic use the set theoretic symbol for an implication relation. I'm partial to the arrow for reasons that will become clear when we introduce the final symbol below, but it's still important to be aware of both symbols in case you come across the other common symbol somewhere.

Another way of thinking of the sentence at the beginning of this subsection is as one stating something like "Following along well so far **implies** a high likelihood of success in the propositional logic unit." Or maybe "**hypothetically**, if you are following along well, then you'd do fine in the propositional logic unit." Or finally perhaps "On the **condition** that you are following along well so far, you'll do fine in the propositional logic unit." These different possibilities illustrate why we might call a statement like the one at the beginning of this subsection an implication, a hypothetical, or a conditional.

If I say "If you eat a peanut butter and jelly sandwich, you won't be hungry anymore" I've said something *conditional* or *hypothetical*. I haven't said "you will eat a PB&J" and I haven't said "you won't be hungry anymore." Both of these claims could easily be false. What I instead said is that ***IF*** you were to eat a PB&J, ***THEN*** you won't be hungry anymore. There's a conditional relationship between them.

Here are some examples:

Example *[Math Processing Error]*

If he's not going to the dance with you,

that implies that he's going with Sheandra!

Example *[Math Processing Error]*

If I took out Stannis' army at Winterfell,

then I'd have no other rivals to my claim as Warden of the North

Example *[Math Processing Error]*

If you don't do your homework, you won't be allowed to go out tonight.

Example *[Math Processing Error]*

You'll do well on the exam only if you study

Example [Math Processing Error]

A necessary condition for being admitted

is submitting a properly filled out application with all required documents

Example [Math Processing Error]

A sufficient condition for eating is

putting food in your mouth, chewing, and swallowing

Symbolizing a conditional is fairly straightforward. Identify the antecedent—often the proposition between "if" and ", then"—and the consequent—often the proposition following ", then". Once you've identified them, symbolized them (probably using a single variable letter, but the antecedent or consequent could always be more complex than that). Then add an *[Math Processing Error]* or a *[Math Processing Error]* and enclosed in parentheses.

Definition: Antecedent and Consequent

The first part of a conditional—what comes before the arrow/horseshoe—is called the **"antecedent"** and the second part is called the **"consequent."**

Let's take our first example: "*If he's not going to the dance with you, that implies that he's going with Sheandra!" Notice how there's a "not" in the first part (the antecedent), so we'll need to capture that when we're symbolizing. Here is the sentence with the logical words taken out and symbolized:*

[Math Processing Error] he's going to the dance with you [Math Processing Error] he's going with Sheandra

If it's not the case that he's going with you, then he's going with Sheandra. Notice that the "that implies that" is an indicator word for a logical implication relation. Once we're sure that the only English that's left is parts of simple propositions, then we can go ahead and symbolize the simple propositions using variable letters:

*[Math Processing Error]*Y *[Math Processing Error]* S

Finally, you'll enclose the whole thing—antecedent, arrow, consequent—in parentheses:

(*[Math Processing Error]*Y *[Math Processing Error]* S)

Be careful not to do this:

[Math Processing Error](Y *[Math Processing Error]* S)

If you don't enclose the negation, you get something very different. The statement immediately above symbolizes a statement like "it's not the case that if he's going to the dance with you, he's going with Sheandra." But that's not what our original statement said. Even worse, "*[Math Processing Error]*(Y *[Math Processing Error]* S)" is logically equivalent to (Y and *[Math Processing Error]*S), which means "he is going to the dance with you and he's not going with Sheandra." But that's definitely not what our original statement said or implied. So the lesson here is to always take care to enclose the *whole antecedent* (or disjunct, or conjunct) when putting the parentheses around an operator like a *[Math Processing Error]*, *[Math Processing Error]*, or *[Math Processing Error]*.

There's a bit more to say about each of these symbols, but for now let us press on. We'll address some of this more in-depth detail in the following section (6.3 More Thoughts on Symbolization.)

"If and only if..." AKA Equivalence

The final logical symbol in the basic set is sometimes called "material equivalence" or "biconditional." Here are some examples:

Example [Math Processing Error]

A person is Commander-in-Chief of the United States military if, but only if they are the President of the United States

Example [Math Processing Error]

When and only when you clean your room will you be given your allowance

Example [Math Processing Error]

If and only if I win the election will I make you my Chief of Staff

Example [Math Processing Error]

Owning a home is a necessary and sufficient condition for being a homeowner

We might not see this one quite as commonly in ordinary language, but it does arise sometimes.

Here're the symbols we use for equivalence or biconditional:

[Math Processing Error]	[Math Processing Error]

The left one is called a "double arrow" and the right one is called a "triple bar". Now you can see why I prefer to use the arrows rather than [Math Processing Error] and [Math Processing Error] : there's a clearer relationship between the symbols and that relationship is reflected in the logic of the symbols. Let's look at an example. The second one from above actually means something like:

If I win the election, I will make you my Chief of Staff;

but only if I win the election, will I make you my Chief of Staff

Which we would symbolize in the following way:

((W[Math Processing Error]C) [Math Processing Error] (C[Math Processing Error]W))

Notice how there's an arrow going W to C and another going C to W? Well, the double arrow does all of this work without something as long and complex as this. We can just write:

(W [Math Processing Error] C)

And it means the exact same thing. These two logical formulas are logically equivalent. It's as if this formula means:

W[Math Processing Error]C

W[Math Processing Error]C

Seeing this relationship is key to understanding what equivalence is as a logical relationship.

Warning, though: "[Math Processing Error]" is NOT a symbol in propositional logic. We only use it to illustrate things in textbooks, never when actually symbolizing or otherwise doing logic.

A final note about "if and only if" before pressing on: sometimes logicians and philosophers will use the word "iff" as shorthand for "if and only if." When you see "iff" think "if and only if" and then think [Math Processing Error] or [Math Processing Error].

Okay, so now that you have a basic introduction to the five logical symbols, we can think a bit more technically about how to symbolize English sentences in propositional logic. Remember: the goal here is to have a sort of diagram of the logical structure of sentences and arguments so that eventually we can manipulate the symbols in that "diagram" to find out more about the logical structure.

Logical Words and Operators Summary

Here is a table of the symbols you'll see and the English logical words and phrases they are meant to symbolize. Note that there are more than one symbol for the same logical operator. This is because some textbooks use different symbols and It's important to be aware of both symbols so you have access to a wider variety of practice problems and other resources (including these textbook chapters).

Table *[Math Processing Error]*

Not, it is not the case that, it is false that	*[Math Processing Error]* (hook), ~ (tilde)
And, yet, but, however, moreover, nevertheless, still, also, although, both, additionally, furthermore, along with, in addition to, ...	*[Math Processing Error]* (wedge), *[Math Processing Error]* (dot), & (ampersand)
Or, unless	*[Math Processing Error]* (vee)
If...then..., only if, implies, given that, in case, provided that, on condition that, sufficient condition for, necessary condition for	*[Math Processing Error]* (arrow), *[Math Processing Error]* (horseshoe)
If and only if, iff, is equivalent to, necessary and sufficient condition for	*[Math Processing Error]* (double arrow), *[Math Processing Error]* (triple bar)

7.3: More Thoughts on Symbolization

The Grammar of Propositional Logic

The five symbols are sometimes called “logical operators”. There’s a reason for this that we’ll get to in section 6.4, but for now be aware that when I say “operator” I am referring to one of the five logical symbols (or equivalent symbol pairs) that we’ve learned so far.

Logic has a grammar or syntax just like English has a grammar. If I was to say “the milked I stars goats those under,” my sentence wouldn’t make any sense—and not because I’ve chosen words that don’t mean anything in that combination. Indeed, if I put them in a *grammatical* order, then suddenly I’ve got a sentence that makes sense (even if it’s a bit strange): “I milked those goats under the stars.” Logic is similarly sensitive to the order we put symbols in—one mistake and suddenly it doesn’t make sense.

Here are the rules of grammar for Propositional Logic—called its “syntax”. One way to think about a syntax is that it’s a set of rules for what counts as a “Well-Formed Formula” or a properly put-together string of symbols. Sometimes logicians call these WFFs (pronounced by logicians “woofs”) or simply “formulas” (after all, we don’t tend to talk about non-well-formed formulas).

1. Any atomic proposition (sentence letter) is a formula.

Example *[Math Processing Error]*

Examples:

- G
- F
- X
- Z

2. Sticking a negation to the left of any formula (any well-formed formula at all, no matter how long or complex) produces a new formula.

Example *[Math Processing Error]*

Examples:

- ~G
- ~(F*[Math Processing Error]*D)
- *[Math Processing Error]*[D*[Math Processing Error]*(R*[Math Processing Error]*(Q*[Math Processing Error]*X))]

3. The other four operators (other than negation) are called “Connectives”. Putting a connective between any two well-formed formulas and then enclosing the whole in parentheses (or brackets) produces a new well-formed formula.

Example *[Math Processing Error]*

Examples:

- (F*[Math Processing Error]*D)
- (~R*[Math Processing Error]*~(Q*[Math Processing Error]*~X))
- [D*[Math Processing Error]*(R*[Math Processing Error]*(Q*[Math Processing Error]*X))]

End of list. These are the only strings of symbols that count as formulas (well-formed formulas, that is): atomic sentence letters, formulas with a negation to the left of them, and connectives with formulas on either side (and parentheses outside of them).

Here are some rules that follow from this definition of the syntax of propositional logic:

1. A Negation is not a connective and a connective must have 2 “arguments” or formulas on either side.

Example [Math Processing Error]

Examples of Non-Formulas:

- D~S
- [D[Math Processing Error](R[Math Processing Error](Q[Math Processing Error]X))]

2. Every *connective* needs a set of parentheses or brackets

Example [Math Processing Error]

Examples of Non-Formulas:

- Q[Math Processing Error]~X
- D[Math Processing Error](R[Math Processing Error](Q[Math Processing Error]X))

3. Never put two formulas adjacent to one another

Example [Math Processing Error]

Examples of Non-Formulas:

- ~G(F[Math Processing Error]D)
- [Math Processing Error][D(R[Math Processing Error](QX))]

4. Never put a negation to the left of a connective or close parenthesis.

Example [Math Processing Error]

Examples of Non-Formulas:

- (F~[Math Processing Error]D)
- [D[Math Processing Error](R[Math Processing Error](Q[Math Processing Error]X))[Math Processing Error]]

To expand a bit on the structure of the operators, it's helpful to note that negation is a *monadic* operator, whereas each connective is a *dyadic* operator. What this means is that a negation takes *one* (monadic as in 'mono' as in "mono-a-mono" or one-on-one) input whereas the connectives or dyadic (dyadic as in 'di' as in two as in "dilemma" or "two options") operators take *two* inputs. So a negation looks like this:

~_____________

Where any well-formed formula at all can go in the blank spot. And the connectives look like this:

(______ [Math Processing Error] ______)

(______ [Math Processing Error] ______)

(______ [Math Processing Error] ______)

(______ [Math Processing Error] ______)

Where, again, any WFF—anything from "**A**" to "**[(~A v ~B)** [Math Processing Error] **~(A** [Math Processing Error] **B)]**"—can go in the blank spots.

Examples follow. Notice all the different sorts of formulas that can go in the blank spots:

Example [Math Processing Error]

[(~A v ~B) [Math Processing Error] ~(A [Math Processing Error] B)]

[(P [Math Processing Error] ~Q) [Math Processing Error] (~P [Math Processing Error] Q)]

[Z [Math Processing Error] ~((D[Math Processing Error]~X)[Math Processing Error](G v H))]

[((G v H) *[Math Processing Error]* ~Q) *[Math Processing Error]* ~D]

It's not hard to see why the dyadic operators are called "connectives": they connect formulas together in different ways.

Tricky Conditionals

It's quite easy to make mistakes in understanding the various conditional propositions and the order in which one should symbolize them. It's worth spending a moment to learn how to properly order conditional propositions.

The first thing to note is that there are essentially three separate conditional phrases that indicate different conditional relationships:

If...(then)...

Only if

If and Only if

Remember now and for the rest of your days that these are different: they mean different things and they have different logical structures. Thus, we must use different logical symbolization patterns to translate them.

"If"

First, if you run into a sentence with an 'if' by itself (no 'only' before it), then the rule is always to *rewrite the sentence in your head* or on paper as an *if...then...* sentence. For example, if you come across the sentence:

People won't help me if I don't have any money

You should first notice that there's a 'not' at the beginning and hidden in "don't". Keep that in mind for later. Second, you should notice that there is an "if" in the middle of the sentence. This means we need to switch up the order of the sentence until it has the structure "if...then...". Here is the result:

If I don't have any money, then people won't help me

Then we make (again, we can do this all in our heads) the negations more explicit:

***If not** I have any money**, then not** people will help me*

Then I'm ready to symbolize:

(*[Math Processing Error]*M*[Math Processing Error][Math Processing Error]*H)

I chose 'M' for "money" and 'H' for "help".

If, alternatively, you come across a sentence like:

If I've been kind to you, then you must return the favor.

You should notice that it's already in "if...then..." format and so doesn't need to be rearranged. Just symbolize as is:

(K*[Math Processing Error]*R)

Another way of putting this rule is that **whatever comes after a solitary 'if' is the antecedent of the conditional, whether that 'if' is at the beginning or middle of the sentence.**

"Only if"

Second, if you come across an "only if", the rule is to (either on paper or in your head) draw an arrow over the phrase. That arrow will be pointing at the consequent. You may need to rearrange the sentence to make sense of it. Here's an example:

Example *[Math Processing Error]*

You'll get into Stanford only if you keep your grades up.

Solution

Draw an arrow over the words "only if":

You'll get into Stanford [Math Processing Error] you keep your grades up.

And if it, like an arrow requires, has a proposition on both sides, then you're golden:

(S*[Math Processing Error]*G)

'S' for 'Stanford' and 'G' for 'grades'. Alternatively, what happens if the 'only if' is at the beginning of the sentence like the following?

Only if you check your oil and tire pressure regularly will you get to keep your car.

Again, just as before, we draw an arrow over "only if":

[Math Processing Error] you check your oil and tire pressure regularly will you get to keep your car.

But this is not like before because there isn't anything to the left of the arrow. This is bad, since arrows need a proposition on either side. Where do we get that other proposition? Well, usually we'll identify the word 'will', which acts a bit like 'then' in an "if...then..." sentence. Then we take what comes after the 'will' and move it to the front—that's our antecedent.

you will get to keep your car [Math Processing Error] you check your oil and tire pressure regularly

Now we're ready to symbolize, but before we do, notice that there's an 'and' in the consequent.

(C*[Math Processing Error]*(O*[Math Processing Error]*T))

'C' for "you get to keep your car", 'O' for "you check your oil regularly", and 'T' for "you check your tire pressure regularly".

"If and Only If"

If we see "if and only if" (or "iff" for short), then we ignore what we've learned about dealing with 'if' and 'only if' separately. Instead, we simply use a double arrow or triple bar (*[Math Processing Error]* or *[Math Processing Error]*). Again, we might find that we need to rearrange the pieces a bit. Here's an example where we need to do some rearranging:

✓ Example *[Math Processing Error]*

If, but only if, you get a job will you be able to afford a new car

Solution

The rule here is to draw a double arrow over whatever phrase means "if and only if". Here, it's the first phrase:

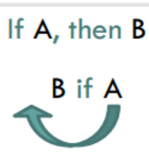
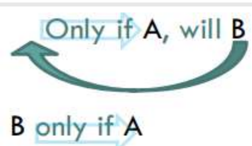
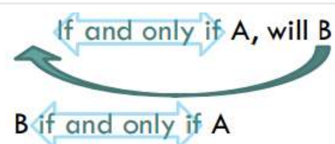
[Math Processing Error], you get a job will you be able to afford a new car

Notice how, again, we've nothing to the left of our connective symbol. But we need something over there. So we, again, look for 'will' and then move the last bit to the front before symbolizing:

you will be able to afford a new car [Math Processing Error], you get a job

(C *[Math Processing Error]* J)

Summary

		
Translates as...	**Translates as...**	**Translates as...**
(A*[Math Processing Error]*B)	(B*[Math Processing Error]*A)	(B*[Math Processing Error]*A)

Why are they symbolized like this? Remember that an arrow means "if the antecedent (left) happens, then the consequent (right) will happen". Let's rehearse some of our examples.

People won't help me if I don't have any money

The claim here is that if this person doesn't have money, then people won't help them. That is, they won't be poor and find people to help them. They'll need money if they want help. Thus the "I don't have money" is the antecedent.

You'll get into Stanford only if you keep your grades up.

This isn't claiming that you *will* get into Stanford if you keep your grades up. Of course not! Stanford is among the most selective schools in the world, so even if you do extremely well, there's still no guarantee. The claim is something more like "if you don't keep your grades up, then you won't get into Stanford." Keeping your grades up is a necessary condition for getting into Stanford. You won't get into Stanford without also keeping your grades up. So if you happen to get into Stanford, it follows that you must have kept your grades up. Hence "you keep your grades up" is the consequent.

If, but only if, you get a job will you be able to afford a new car

This is essentially claiming two things: if you get a job, you will in fact be able to afford a new car; and if you can afford a new car, you must have gotten a job. It's a biconditional or equivalence.

Necessary and Sufficient Conditions

Translating from statements about Necessary and Sufficient Conditions is relatively straightforward. *Necessary conditions go to the right of an arrow*. **Sufficient conditions go to the left of an arrow. Necessary *and* Sufficient Conditions get a double arrow/triple bar**. Here are some examples:

Example *[Math Processing Error]*

If you can turn in a completed term paper by the end of finals week, that'd be sufficient to pass the class.

Solution

What's the sufficient condition? Right-o, "turning in a completed term paper by the end of finals week". Where does the sufficient condition go? Right-o again! To the left of the arrow:

Turn in completed paper by the end of finals week [Math Processing Error] pass the class

(C *[Math Processing Error]* P)

'C' for "Turn in completed paper by the end of finals week" and 'P' for "pass the class". Here's another example:

Example *[Math Processing Error]*

Getting an A in participation is necessary for getting an A in the class

Solution

What's the necessary condition? Righty-O-Daniels, "getting an A in participation". And where do necessary conditions go? Righty Righty Mighty Tidy! To the right of the arrow:

getting an A in the class [Math Processing Error] Getting an A in participation

(C *[Math Processing Error]* P)

'C' for "getting an A in the class" and 'P' for "getting an A in participation." Finally:

Being President is necessary and sufficient for being Commander-in-Chief

What's the necessary and sufficient condition? Being President. Where does the condition go?

Really, it doesn't much matter since (P*[Math Processing Error]*C) is logically equivalent to (C*[Math Processing Error]*P). Mostly we just keep these statements in the order in which they are written:

(P*[Math Processing Error]*C)

'P' for "Being President" or "One is president", and 'C" for "being Commander-in-Chief" or "one is Commander-in-Chief".

Tricky Ands, Nots, and Nors

I can't go to the dance with both you and your friend. Neither you nor your friend are welcome at my house. Either you or your friend won't be able to go to the dance because you'll have to stay here and clean up this mess. I'm not going to the dance with you and I'm not going to the dance with your friend. Each of these statements has a different logical form, but some of them have a special logical relationship with one another. Let's take similar statements and think about how to translate them.

1. Neither Baixa (pronounced Bye-ee-sha) nor Ramón will be president of the club next year.

- We translate "neither-nor" statements as "not or":
- *[Math Processing Error]*(B *[Math Processing Error]* R)

2. I'm not going to fire either Xia or Bo.

- We translate "either not A or not B" statements as:
- (*[Math Processing Error]*X *[Math Processing Error]* *[Math Processing Error]*B)

3. Nia and Goran can't both go in our car with us.

- We translate "not both" statements as:
- *[Math Processing Error]*(N *[Math Processing Error]* G)

4. You can't be here and your friend can't be here.

- We translate "not A and not B" statements as:
- (*[Math Processing Error]*Y *[Math Processing Error]* *[Math Processing Error]*F)

If I tell you that you can't date Tamik and you can't date Peter, then it follows that you can date *neither* Tamik *nor* Peter. Can you see how the following two formulas are logically equivalent?

(*[Math Processing Error]*T *[Math Processing Error]* *[Math Processing Error]*P)

[Math Processing Error](T *[Math Processing Error]* P)

Similarly, notice how the following makes perfect sense: "You'll have to choose which to not have tonight: no ice cream or no cookies—you can't have *both* ice cream *and* cookies." So these two are logically equivalent:

(*[Math Processing Error]*I *[Math Processing Error]* *[Math Processing Error]*C)

[Math Processing Error](I *[Math Processing Error]* C)

These equivalences have a special name: De Morgan's Theorems. They're named after a logician named Augustus De Morgan. We'll go over this in a bit more depth when we get to Natural Deduction in Chapter 6.

For now, keep in mind that these are equivalent, but *always translate in the way that best captures the ordinary language sentence*. Don't translate using the alternative equivalent forms. Here are some more examples:

- Either you won't pass the class or you won't fail the midterm.
  - (*[Math Processing Error]*P *[Math Processing Error]* *[Math Processing Error]*F)
- Neither of these dogs is going to work for me.
  - *[Math Processing Error]*(A *[Math Processing Error]* B)
  - *Notice how I have to capture the sense in which this statement is about neither "dog A" nor "dog b", even though it doesn't say as much. We have to capture the fact that it's a "neither-nor" statement even though the disjuncts don't quite show up in the English sentence.*
- Neither you nor I will be graduating on time.
  - *[Math Processing Error]*(Y *[Math Processing Error]* I)
- We're not both going to be able to go out tonight.
  - *[Math Processing Error]*(Y *[Math Processing Error]* I)
  - *Again, the disjuncts "you are able to go out tonight" and "I am able to go out tonight" aren't really explicit in the English sentence, but we still have to capture the fact that this is a "not both" statement, and so we have to fill in the subjects "you" and "I".*
- That window isn't clean and that other one over there isn't clean.
  - (*[Math Processing Error]*T *[Math Processing Error]* *[Math Processing Error]*O)
  - I went with T and O for "That window" and "the Other window".

Exclusive vs. Inclusive "Or"

When you’re on an airplane and the attendant comes by, asking “would you like cream or sugar in your coffee?”, it’s okay to say “yes, I’d like both, please.” It would be odd if the attendant said, “I’m sorry, I said would you like cream *OR* sugar, I can’t give you both.”

Alternatively, if you’re voting, you often get the choice between a variety of candidates. Let’s simplify it to a choice between 2 candidates. You can vote for either one or the other candidate, but you cannot vote for both. If you do, your vote won’t count at all.

These are both cases where “or” is used, but the first one is what we might call “inclusive” in that both are allowed. Alternatively, the second one is called “exclusive” in that only one or the other is allowed.

The way our “or” operator in standard propositional logic works is as an *inclusive or*: a disjunction is true if both of its disjuncts end up being true. Here’s an example:

Example *[Math Processing Error]*

Either the maid or the butler is the killer

Solution

If we interpret this *exclusively*, then the killer cannot be *both* the maid and the butler. It must be one or the other. If we interpret this *inclusively*, then the killer could be the maid, it could be the butler, *or* it could be both of them conspiring. Central point: there are two ways of understanding an English ‘or’, and the propositional logic symbol ‘*[Math Processing Error]*’ stands for the *inclusive* or.

There are two ways of understanding an English ‘or’, and the propositional logic symbol ‘*[Math Processing Error]*’ stands for the *inclusive* or.

Given that ‘*[Math Processing Error]*’ means inclusive or, how would we go about translating *exclusive or* or XOR? Well, one answer would be to simply invent a new symbol. Some folks already have. Here are a few different options:

[Math Processing Error]	*[Math Processing Error]*	⊻	*[Math Processing Error]*

In this text, we won’t use these symbols. We’ll stick to our original five symbols instead. In principle you could use as many symbols as you like, but it makes certain things harder, so we’ll just stick with our five.

Using just our five symbols, though, we can translate an exclusive or. How might we do that? Think through it step-by-step. An exclusive or is saying something like:

Either A or B will happen, but not both A and B

How do we translate “Either A or B”? Simple:

(A *[Math Processing Error]* B)

‘but’ becomes ‘*[Math Processing Error]*’, but how might we translate “not both A and B”? Again, we’ll go step-by-step. How do we translate “both A and B”? Simple:

(A *[Math Processing Error]* B)

And since it’s “not both,” we simply put a *[Math Processing Error]* in front of it:

[Math Processing Error](A *[Math Processing Error]* B)

So when we stick all of this together, we get:

(A *[Math Processing Error]* B) *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* B)

Either A or B, but not both A and B. Oh! And don’t forget the brackets that go with that middle ‘*[Math Processing Error]*’:

[(A *[Math Processing Error]* B) *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* B)]

As a rule, when you see an ‘or’ in a practice problem, the way to symbolize it will be with a simple ‘*[Math Processing Error]*’, but occasionally it’s clearly meant to communicate the logical content of the exclusive or, so you’ll need to use this formula above.

Framing Words

English grammar has some really helpful features that reveal the logical form of sentences. These are what we might call "framing words." Here are a few English sentences to look at:

Either you hand over the money or this will end in violence

If you both pass your exams and you get a summer job, then we'll help you out with buying your new car.

Neither will you pass go nor will you collect $200.

Each of these sentences has its logical form built into it's grammatical structure in a fairly explicit way. Let's explore each in turn.

The first sentence has an "Either...or..." structure. When we see this, we can almost always treat it this way:

Either __[One disjunct]___, Or ___[the Other disjunct]___

So our first sentence:

Either you hand over the money or this will end in violence

Will be interpreted in the following way:

[Math Processing Error]

We simply take whatever comes between the "either" and the "or" and call that one disjunct, and then we take whatever comes after the "or" and call that the other disjunct—*even if they are logically complex*. That is, even if there are logical words between the "either" and the "or", we can still count all of that as one disjunct. Here's another example to illustrate this latest point:

Either you both hand over the money and don't call the police, or this will end in violence.

Notice how there's a "both...and..." between the 'either' and the 'or'? That's okay, we'll just treat the 'both...and' as, you guessed it, a conjunction; and then we'll treat that whole conjunction as one disjunct of the overall disjunction. Here's how it looks:

[Math Processing Error]([you hand over the money] AND [you don't call the police]) OR

[this will end in violence][Math Processing Error]

So our final logical formula looks like this:

((M *[Math Processing Error]* ~P) *[Math Processing Error]* V)

And keep in mind that because we have equivalent symbol pairs, here is an equivalent formula that means *exactly* the same thing:

((M *[Math Processing Error]* *[Math Processing Error]*P) *[Math Processing Error]* V)

Okay, now let's look at our second example from above:

If you both pass your exams and you get a summer job, then we'll help you out with buying your new car.

[Math Processing Error]([you pass your exams] AND [you get a summer job]) IMPLIES

[we will help you buy your new car][Math Processing Error]

And the final logical formula looks like this:

((E *[Math Processing Error]* J) *[Math Processing Error]* C)

Let's look at our final example really quickly:

Neither will you pass go nor will you collect $200.

"neither...nor" works essentially the same way as "either or" except that it's negated, so whatever comes between 'neither' and 'nor' is one disjunct and whatever follows 'or' is another disjunct. Here is how this particular example works out:

NEITHER ([you will pass go] NOR [you will collect $200])

Here's the final logical formula:

[Math Processing Error](G *[Math Processing Error]* C)

What are the general rules that allow us to make full use of framing words? You're in luck, I've compiled at least a great deal of them. Here they are:

Only if [consequent]**, will** [antecedent]. [antecedent] **only if** [consequent].
[equivalent] **if and only if** [equivalent]. **If and only if** [equivalent]**, will** [equivalent].
If [antecedent]**, then** [consequent]. [consequent] **if** [antecedent].
Both [conjunct]**, and** [conjunct]. • Not both___, and ____
Either [disjunct]**, or** [disjunct]. • Neither [disjunct], nor [disjunct]

We can use these framing words to translate even *very* complex sentences from ordinary language into a propositional logic formula.

Here's an example:

Example [Math Processing Error]

If and only if a person is both arrested at the border and is carrying narcotics, will they be either both deported and have their visa revoked if they are not a citizen or both incarcerated and prosecuted if they are a citizen.

Solution

Woah! I need to take a nap!

Take your nap if you need it, but we do have all of the tools necessary to make sense of what's going on here. Remember that we are just using symbols to represent the logical form of the ordinary language sentences here, and you already have some intuitive understanding of the logical form because you are an English speaker, so we're just trying to capture what you already intuitively know. As always, if we go step-by-step we won't get overwhelmed.

First: what's the overall grammatical structure of the sentence? In other words: what are the framing words that set up the frame for the whole sentence?

Well, it starts with "if", so "if...then"?

Remember, dear student, that "if and only if" is a different thing from "if", so we have to learn to identify them and treat them differently.

Oh, so "if and only if...will?"

Yeppers. Good work. So we already know that the main operator of the sentence—the logical operator that belongs to the outer most parentheses—is... what?

A double arrow thingy

Yep, or a triple bar (*[Math Processing Error]* or *[Math Processing Error]*). So the overall structure is like the following:

(??? *[Math Processing Error]* ???)

What goes in place of those question marks? Well, the first thing to find out is where each equivalent starts and stops. We look between the "if and only if" and the ", will" for one equivalent:

a person is both arrested at the border and is carrying narcotics

And then we look after the ", will" for the other equivalent:

they be either both deported and have their visa revoked if they are not a citizen or both incarcerated and prosecuted if they are a citizen.

Now we can just go equivalent-by-equivalent. Let's start with the first one—it's shorter. We see a "both...and..." here, which tells us we'll be working with a...what?

Conjunction!

Yes indeedy. Well done! So that's fairly easy to do. The left equivalent will be a simple conjunction, which means the whole formula as far as we've gotten will look like this:

((A *[Math Processing Error]* N) *[Math Processing Error]* ???)

Now for the right equivalent, which is quite complicated. Here I've put the logical words in bold:

*they be **either both** deported **and** have their visa revoked **if** they are not a citizen **or both** incarcerated **and** prosecuted **if** they are a citizen.*

Again, we'll just work step-by-step. We see and "either", so that means we'll be looking for our "or". What comes between those two is one *disjunct* or half of a disjunction. Here's that disjunct:

***both** deported **and** have their visa revoked **if** they are not a citizen*

To reorient, we're working on the left disjunct of the right equivalent, here:

((A *[Math Processing Error]* N) *[Math Processing Error]* (*[Math Processing Error]* *[Math Processing Error]* ???))

Okay, what do we see there? A "both...and...", which tells us that we're again working with a conjunction. What else? An "if", which tells us we'll have a conditional to sort out. Remember that when we see an 'if' in the middle of a sentence, we need to mentally rewrite the sentence as an "if...then..." sentence. Technically, there's a bit of ambiguity in the grammar here. I'm just going to default to the most likely reading, which is "if they are not a citizen, then they will both be deported and have their visa revoked." So our left disjunct looks like this:

((A *[Math Processing Error]* N) *[Math Processing Error]* (*[Math Processing Error]* *[Math Processing Error]* ???))

Now we need to focus on that last disjunct where the question marks are. That's the part that comes after the final 'or' in our example. It reads:

***both** incarcerated **and** prosecuted **if** they are a citizen*

This will work essentially the same as the last disjunct. Rewrite as an "if..then" and then note that the consequent is a "both...and..." (a conjunction). Then symbolize:

((A *[Math Processing Error]* N) *[Math Processing Error]* (*[Math Processing Error]* *[Math Processing Error]* (C *[Math Processing Error]*(I *[Math Processing Error]* P))))

Every time I choose a new Sentence Letter, I'm always checking the rest of the formula to ensure that I'm not reusing a letter that actually signifies a different proposition. I am sure to reuse that 'C', though, since there's a clear logical relation between "x is a citizen" and "x is not a citizen"—one is a negation of the other and so our logical has to capture this fact.

That's it. We're all done! Good job! That was tough, but you can see how when we go step-by-step, we're able to do this accurately and without much headache.

On an unrelated note, some cases use the same framing words as we have learned, but aren't as clear. Here are a few examples:

Example *[Math Processing Error]*

Both you and your sister will need to clean your rooms before you can go out with friends tonight.

Example *[Math Processing Error]*

Either of you can drive the car today, but not your friends.

Think about what each is saying and then try your best to capture the idea in a symbolization. For instance, the first one seems to be saying something like:

Both you will need to clean your room before you can go out with friends tonight, and your sister will need to clean her room before she can go out with friends tonight.

There is a sense in which what is being said here is somewhat different than I have written above. It might mean something like:

Only if both of you clean your rooms will both of you get to go out with your friends tonight.

Which might break down into:

Only if you clean your room will you get to go out with your friends tonight AND only if your sister cleans her room will she get to go out with her friends tonight.

This might capture what's being said fairly well, but we'll stick to doing something a bit less sophisticated: we'll stick to trying to capture the surface content of the English sentences and so we'll focus only on explicit logical indicator words. So if we stick with the original reformulation, we get something with a logical structure like this:

[you will need to clean your room before you can go out with friends tonight]

AND

[your sister will need to clean her room before she can go out with friends tonight.]

It's a simple conjunction: the bracketed sentences don't have any internal logic to them—they're just descriptions of the world, true or false. With that in mind, the symbolization is fairly straightforward:

(Y *[Math Processing Error]* S)

'Y' for "you will need to clean your room before you can go out with friends tonight" and 'S' for "your sister will need to clean her room before she can go out with friends tonight."

Now let's take the second example from above:

Either of you can drive the car today, but not your friends.

This has an "Either...Or..." structure to it, but the actual disjunct propositions are hidden in the grammar a bit. If you think through it, you'll find that the first part (before ", but") is a disjunction between:

You can drive the car today ***OR*** *(the other) You can drive the car today*

This sentence is grammatically strange for our purposes in a few ways: it's second person plural, meaning it's a statement to two "you"s, and the two disjuncts are a bit hidden. It has an 'either' that isn't followed up with an 'or', and so the disjunction isn't as clear as it is when you have that nice frame to work with.

Now that we've got the hard part figured out, we can go ahead and symbolize the whole sentence:

[(Y *[Math Processing Error]* O) *[Math Processing Error]* *[Math Processing Error]*F]

You can drive the car today **OR** (the other) You can drive the car today, **BUT** your friends can **NOT** drive the car today.

7.4: Logical Operators as Truth Functions

What's a Truth Function?

As we delve a bit deeper in learning about our five logical operators, the next step is to explore exactly what they mean and how they work. We might have an intuitive sense of what they mean, but here we'll solidify that understanding in a logically precise way.

If I tell you Barack Obama is president, then I've told you that the world is one determinate way. Alternatively, if I tell you that

Either *Barack Obama is president* Or *he's retired*

then I've told you that we live in one of two possible worlds: one world where Obama is president right now or another world where Obama is retired right now. Which world do we actually live in?

Notice how this *disjunctive* statement is true or false depending on the truth or falsity of its *parts*. A fancy logician's way of saying essentially the same thing is that the truth of the whole statement is a *function* of the truth of its parts. This statement is very clearly dependent on its parts:

I will both graduate high school and get into college

The statement as a whole is false if the person never graduates high school (even if they get into college!). It's also a false statement is the person never gets into college (even if they graduate high school!). The statement as a whole is only going to be true if *both* of the conjuncts end up being true.

This is what it means to say that our logical operators ($\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$) are *Truth Functions* or *Truth Functional Operators*—they are little machines that tell us whether a whole statement is true based on whether the component parts are true.

Think about this intuitively for a second:

If I tell you that I'm either going to get rich or I'm going to get arrested, my statement will be true if I get rich, it will be true if I get arrested, and it will probably also count as true if I get rich *and get arrested!* It's a disjunction, and disjunctions are true if just one of their disjuncts are true.

If I tell you that "I am not getting into UCLA," but then I *do* get into UCLA, I have said something false, right? A negation is true if the negated statement is false, and false if the negated statement is true. If I say "I am not going to that party tonight," and then I end up not showing up, then I have said something *true*: the proposition "I am going to that part tonight" ended up being false and so the negated version of it ends up being true.

If I tell you that "you'll get your allowance if and only if you do your chores," but then you do your chores and don't get your allowance, then my statement will have been false: turns out doing your chores wasn't perfectly sufficient for getting your allowance.

Each operator, in other words, takes certain inputs (the truth or falsity of the atomic sentences that make up its disjuncts, conjuncts, antecedent and consequent, or the like) and delivers one output: a true or a false. This is a logical operation. It's sort of like each of our five logical operators are little machines.

Think about it like a vending machine:

I need to input the correct set of inputs to get the output I want. This is because **every set of inputs corresponds to a particular output**. If I put money in and then push “D-7”, I’ll get a soda, but if I put a different amount of money in and then push “G-8”, I’ll get a candy bar. Different inputs means a different output.

Now for Logical Operators, we only have two possible outputs: **True** and **False**. These two values are called “**truth values**”. It’s a bit confusing because only one of them is “true”, but that’s what they’re called. If it’s helpful, think of them like numerical values. If I am writing a number in the standard decimal system, then I can fill in values 0-9 for the 1’s place, the 10’s place, the 100’s place, and so on. 0-9 are the *values* that can go into those places. For instance, “106” has a ‘1’ in the hundreds place, a ‘0’ in the tens, and a ‘6’ in the ones places. Instead of ten different values, though, logic only works with two values: ‘true’ and ‘false.’ It’s kind of like binary, where ‘0’ and ‘1’ are the only values you have to work with.

If each operator gives an *output* of “true” or “false”, what sorts of *inputs* do we give into the system? Well, the inputs are *more truth values*. A conjunction takes two T’s and turns them into one output of T. A disjunction takes an input of T and F and outputs T. A negation takes an input of T and outputs F. Each operator takes some set of truth values as an input (1 or 2 truth values) and then outputs just one truth value. Note: negation is the only operator that takes just one input. The other four take two inputs.

A function is something that takes an input and gives you an output. So logical operators are **Truth Functions** in the sense that they take truth values as inputs and give you truth values as an output.

Truth functions work together in a particular order to determine what individual output will result from a set of inputs. So just as:

$$\neg \text{A}$$

Is a truth function or a truth functional formula, so is:

$$\neg(\text{A} \leftrightarrow \text{B})$$

And also

$$[\neg[\text{D} \leftrightarrow \neg(\text{X} \rightarrow (\text{Z} \bullet \text{Q}))] \vee \text{P}]$$

It ain’t pretty, but it *is* a truth functional formula in that given a set of inputs (one input for each sentence letter), it will give exactly one output. The parentheses and brackets determine the order in which order we apply the operations, but we’ll get into that later. For now, just keep in mind that a set of inputs will give a single output for any WFF in propositional logic:

$$\begin{array}{l} D : False \\ X : True \\ Z : True \\ Q : False \\ P : True \end{array} \Rightarrow [\neg[D \leftrightarrow \neg(X \rightarrow (Z \bullet Q))] \vee P] \Rightarrow True$$

What’s a Truth Table?

A truth table is just that: a table of truth values. It tells you what output you’ll see given some set of inputs. Let’s look at a very simple one so we can see what the different parts are:

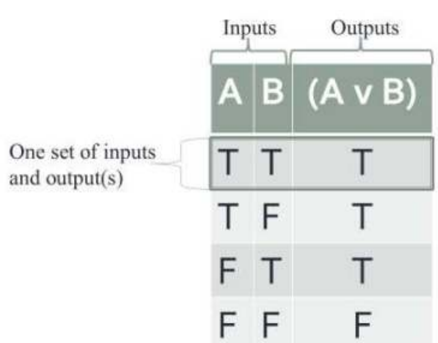


Figure 7.4.1: Anatomy of a truth table

The left side of the truth table is called the “input side” because it houses the inputs that will determine what outputs we find on the right side or “output side”. A complete set of inputs and outputs is represented in each **row** of the truth table. The boxed row above has two T's as inputs and a T as an output. Remember that there is always just one output in classical propositional logic, even though there might be lots of inputs.

The inputs are the truth values of the atomic sentence letters that make up part of the formula. The formula(s) that define(s) the truth table are found on the top right heading above the outputs.

So what this table is telling us is that if A and B are both Truth, then (A ∨ B) will be true. Similarly if A is True and B is False (second row) or if A is False and B is True (third row). But if (fourth row) A and B are both False, then (A ∨ B) will be false as well.

Definitions of Logical Operators

The following table summarizes the "truth functional profile" of each operator: it tells you what the output will be given various inputs.

A	B	~B	(A•B)	(A∨B)	(A→B)	(A↔B)
BT	T	F	**T**	T	T	**T**
BT	F	T	F	T	**F**	F
BF	T		F	T	T	F
BF	F		F	**F**	T	**T**

How do we read this table? It's easier than it looks. That thick vertical white line between "B" and "~B" divides the inputs from the outputs. So the "A" and "B" columns are the inputs columns and the rest are telling us what the output of various operators will be. The simplest column is "~B": it tells us that when the proposition being negated is True (first row where B is "T"), ~B will be False, and when the proposition being negated is False (second row where B is "F"), ~B will be True.

The other columns are a bit trickier since they involve the inputs of a truth value for B *and* a truth value for A. The (A•B) column is telling us that (A•B) is true only when both inputs are "T"s. Otherwise, it's false.

I've highlighted the "odd one out" for each column. This is the easiest way to memorize this information: figure out what the exception to the rule is and memorize that. (A•B) is false for all but one set of inputs (T and T). (A∨B) is true for all but one set of inputs (F and F). (A→B) is true for all but one set of inputs (T and F, in that order). Finally, (A↔B) is true when the inputs are the same and false otherwise.

Here's a different way of representing the same information:

- ~A is the opposite of A
- (A•B) is true only when both are true
- (AvB) is false only when both are false
- (A⊃B) is false only when A is T and B is F
- (A≡B) is true only when A and B are the same

Computing Truth Values

Now that we know what each truth functional operator *does*, we can perform an operation: we can finally *do* something with logic.

How do we do it? It's simple: we start from the inside and work our way out.

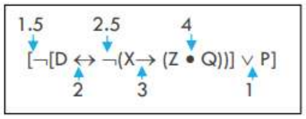
Here are the steps:

Step One: Find the Innermost Operator

The parentheses and brackets enclose the operators. The more parentheses or brackets there are enclosing an operator, the further "inside" the formula the operator is. Our first task is to find the operator that is the most "inside" the formula. Let's practice on this formula:

$$[\neg[D \leftrightarrow \neg(X \rightarrow (Z \bullet Q))] \vee P]$$

A fast way of doing this would be to count up how many *sets* of parentheses or brackets enclose each operator. For negations, we'll pretend that there's a 1/2 set of parentheses around them so we can be sure to get the order correct. I've counted them up here:



Negations don't technically have parentheses or brackets around them, but they also must be calculated *before* the operators that have the same amount of parentheses as they do. For instance, the ∨ and the leftmost ¬ are both inside of only one set of parentheses/brackets, but we are supposed to calculate the value of the negation *before* we calculate the disjunction. So we add an extra .5 as if there was an extra set of parentheses around the negation and what it negates.

What's the innermost operator? It looks like that ● is the innermost: it is enclosed by 4 sets of parentheses/brackets. The → is next, then the ¬ to the left of it. Then the ↔, the leftmost ¬, and finally the **Main Operator**: the ∨.

The Main Operator is the operator that is enclosed by the outermost parentheses/brackets (or, if it's a negation, it's not enclosed at all). The truth value output of the main operator is the truth value of the whole proposition.

Step Two: Fill in the given truth values

We are given the truth values for each atomic sentence. For instance "Barack Obama is President of Canada" is a false proposition, while "Barack Obama was President of the United States" is a true proposition. When we're calculating these truth values, though, we are doing something that is "pure logic" in that it's divorced from the actual truth or falsity of the propositions. At this point, let's just pretend that these truth values come from a random flip of a coin. We're trying to calculate what the truth value of the whole proposition will be given a random set of atomic truth values.

If you're given a practice problem for this section, it will have truth values for each sentence letter given as part of the problem. Here is an example:

[¬[D ↔ ¬(X→ (Z ● Q))] ∨ P]

D, Z, and Q are True

X and P are False

We then, for step two, fill in these truth values like so:

[¬[**T** ↔ ¬(**F**→ (**T** ● **T**))] ∨ **F**]

Now we're ready to start calculating!

Step Three: Calculate from the "Inside-Out"

Now that we've identified the order of the operators, we can calculate the truth values from the inside out.

Each step here is as simple as looking up in the truth table from Definitions of Logical Operators the inputs you're given and then writing down the output from the table. If you understanding the operators a bit more clearly, then you can skip the truth table and go straight to understanding what the output should be. We'll do the painstaking version here just this once.

We decided in step one that the order is: ●, →, the inner ¬, ↔, the leftmost ¬, and finally ∨.

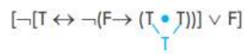
What surrounds the ●?

(T ● T)

What is the truth table for ●?

A	B	(A●B)
T B	**T**	**T**
T **B**	F	F
F **B**	T	F
F **B**	F	F

I've highlighted the relevant row already. The two inputs are T and T, so we want to look at the top row. What's the output? T! We can now move forward in our calculation:



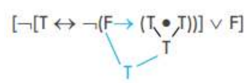
Now to calculate the →. What's the input here? We take the F from before the arrow and the T that we calculated in our previous step. So it's:

F→T

Again, we consult the truth table:

A	B	(A→B)
T **B**	T	T
T **B**	F	F
F **B**	T	T
F **B**	F	T

Looks like the row we're interested in is the 3rd row: F→T. What's the output? T! That T is the truth value of everything inside the parentheses there, so it's the truth value of (F→ (T • T)). Again, we can mark this on our formula:

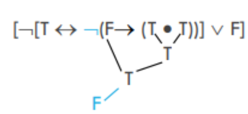


What operator is next? The negation just to the left of that arrow. In this case the negation negates everything in those parentheses to its right (F→ (T • T)), which means it negates that T we just calculated.

What's the truth table for Negation?

A	~A
T	F
F	T

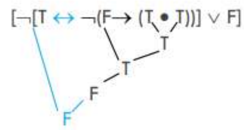
If we're calculating ¬**T**, then we check to see what the output will be when the input is T. So in the table above, when A is T, ~A is F. That means ¬**(F→ (T • T))** is False! Again, we update our formula:



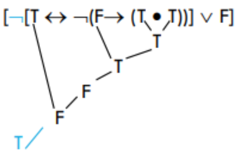
We can now calculate the truth value of the ↔. What are it's inputs? It has T to the left and to the right? ... an F. Good job. What is T↔F? Look it up in the truth table:

A	B	(A↔B)
T **B**	T	T
T **B**	F	F
F **B**	T	F
F **B**	F	T

A "↔" is true only when both sides are the same. Are both sides the same in T↔F? No, of course not. It follows that T↔F is False. Again, update our formula:



Almost done! Calculate the truth value of the negation. Remember: negation just flips whatever it negated. A T to an F and an F to a T. So what would ¬F be? Yep! True.



Last step!!!!! The ∨ connects disjunct **¬[T ↔ ¬(F→ (T • T))]** with disjunct **F**. We now know that the truth value of **¬[T ↔ ¬(F → (T • T))]** is True. What is T ∨ F? Pull up the Truth Table:

A	B	(A∨B)
T **B**	T	T
T **B**	F	T
F **B**	T	T
F **B**	F	F

T ∨ F is the second row here: the output is True. Waaaay back in step one we decided that the main operator was the ∨, so that means we're done once we've calculated the truth value of the ∨. We just did! It's True.

What did we find out? We found out that **[¬[D ↔ ¬(X→ (Z • Q))] ∨ P]** is True in the case that Propositions D, Z, and Q are True and Propositions X and P are False. The more we understand the definitions of the logical operators discussed in 6.4.3, the faster we'll become at calculating complex truth values. Pretty soon it will become second nature.

7.5: Logical Analysis using Truth Tables

Truth tables, which we introduced in the last section, are primarily useful in a different way. Instead of just using them to tell us what the different logical operators mean, we can use them to do some in depth analysis of the logical form of a statement, a set of statements, or an inference.

What's a logical analysis? A logical analysis is a set of things we can do to learn something about a particular logical structure. The statement "I'll go if you don't go or if we can get a babysitter and it's not too expensive for both of us to go" has a particular logical form—something like

[(B*[Math Processing Error]***~E)***[Math Processing Error]***I]**. If we do a logical analysis of this logical form, we'll find out certain things. For instance: when is this statement true? That is, what must the world be like for this statement to turn out to be true? Is this statement always true? Is it always false? What is its relationship with other similar logical forms like **[~(B***[Math Processing Error]***E)***[Math Processing Error]***I]**? **[~(B***[Math Processing Error]***~E)***[Math Processing Error]***~I]**? **[(B***[Math Processing Error]***E)***[Math Processing Error]***~I]**? Is it possible that all of these statements could be true at the same time? That is, are they consistent? We can find out the answers to all of these questions using Truth Tables.

Conceptually speaking, we're doing the following when we build a truth table:

1. Collecting all of the possible combinations of truth values and listing them out.

- *That is, we're finding out all of the different ways that T and F can be combined for our atomic sentence letters.*

2. Using each set of possible truth values to calculate the output truth value for a complex formula (or a set of complex formulas).

- *That is, we're doing what we did in Computing Truth Values: we're taking the truth values of atomic sentence letters as an input and calculating the single truth value output.*

3. Analyzing the results.

- *That is, we're looking at the resulting output and trying to figure out what it tells us about the logical formulas in question.*

The goal is to see what possible conditions of the world (what combinations of true and false for the atomic propositions, each of which either describes the world correctly or incorrectly) give what sorts of truth values for the complex formulas we're analyzing and then to look for patterns in the outputs to tell us something about the individual statement, the set of statements, or the argument we're analyzing.

That's not terribly illuminating, though, about *how to actually go about building* a truth table. Let's look more concretely at what to do.

First, I must point out a new notation that will be unfamiliar to you. Well, you will have seen it before lots of times, but not here in logic. The forward slash!

If you see this:

[*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* P] / (X*[Math Processing Error]* (Z *[Math Processing Error]* Q))]

That means there are two formulas here:

[*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* P]

And

(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))]

The forward slash (/) has no logical meaning. It simply separates formulas from one another so we can list them on a single line without it seeming like they are parts of the same formula.

Building a Truth Table

Building a truth table is very straightforward, but that doesn't mean it's not going to take some getting used to. Let's divide it into a series of steps.

Step One: Figure out what size you need

How many *unique* sentence letters are in the formula or set of formulas? Just count 'em up, counting each unique letter only once (so two B's just count as one). Here are some examples:

- [~(B*[Math Processing Error]*E)*[Math Processing Error]*I] has 3 unique letters: B, E, and I
- [~(B*[Math Processing Error]*E)*[Math Processing Error]*B] has 2 unique letter: B and E
- [~(I*[Math Processing Error]*I)*[Math Processing Error]*I] has 1 unique letter: I
- [~(B*[Math Processing Error]*E)*[Math Processing Error]*I] / [(B*[Math Processing Error]*E)*[Math Processing Error]*~I] have 3 unique letters all together: B, E, and I
- [*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* P] / Z / (P *[Math Processing Error]* Q) has 5 unique letters: D, X, Z, Q, and P

Once you've figure out this magic number, you plug it into a magic formula: 2^n, where n is the magic number: the number of unique sentence letters in the set of propositions the truth table is for. The result of this mathematical formula is the number of rows you'll need in your truth table.

Here are a set of truth table sizes:

# of unique letters	# of rows
1	2
2	4
3	8
4	16
5	32
6	64
16	65,536

Notice the pattern? The nice thing is that chances are, your instructor will only assign at most a 4-letter truth table, so 16 rows is the absolute most you'll usually need to work with. Most instructors stick to 1, 2, and 3-letter truth tables.

The downside of truth tables is that it doesn't take long before you have to start making tables with 32, 64, 128, 256, 512, 65,536 rows! That's too much to really make making a truth table worthwhile. This is called the problem of **Combinatorial Explosion** because all of the combinations that are possible "explode" to astronomical numbers**.** Nevertheless, they are quite useful for relatively simple problems.

Step Two: Make a truth table

A truth table is a table with a column for each unique sentence letter (usually in the order in which they show up in the formulas you are analyzing) and then a column for each formula you are analyzing. Here are some examples:

R	**R***[Math Processing Error]***(R***[Math Processing Error]***R)**	**R***[Math Processing Error][Math Processing Error]***R**

Q	**R**	**R***[Math Processing Error]***(Q***[Math Processing Error]***R)**	**R***[Math Processing Error][Math Processing Error]***Q**

Q	R	R*[Math Processing Error]*(Q*[Math Processing Error]*R)	R*[Math Processing Error][Math Processing Error]*Q

Notice how each column gets its own label and there's a bigger border separating the individual sentence letters (the inputs) from the complex formulas (the outputs). So far so good.

Step Three: Fill in the input side

This step is always the same no matter what the columns labels are. There are basically 3 different tables you'll make, and for bigger tables you can extrapolate the same basic pattern. Remember that the goal when filling in the input side is to make a list of all of the possible combinations of the two truth values T and F.

So when we only have one unique sentence letter. The two possibilities are that the letter is True and the letter is False:

R	R *[Math Processing Error]*(R*[Math Processing Error]*R)	R*[Math Processing Error][Math Processing Error]*R
T		
F		

It's a bit more complicated if we have two unique sentence letters:

Q	R	R*[Math Processing Error]*(Q*[Math Processing Error]*R)	R*[Math Processing Error][Math Processing Error]*Q
R T	T		
R T	F		
R F	T		
R F	F		

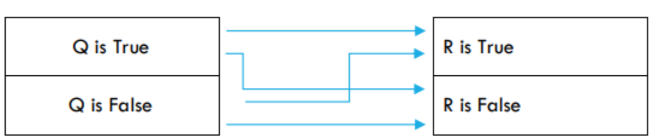
Each row in the table is a possible set of truth values. The possible combos are TT, TF, FT, and FF. See how that works?

Before moving onto an eight row truth table, let's think about the pattern here. We've started on the right nearest the thick line separating the inputs from the outputs and alternated going down: TFTF. For the first table, we alternated too, but just had to do it once: TF. On the second table, we alternated TF going down the R column, and then just repeated so the downward pattern would be TFTF.

Then we moved left and alternated every 2, so the Q column reads TTFF. Why would we do this? Well, we have R is true and R is false, then we need to test for when Q is True with both of these possibilities, and then test again for when Q is false. The result is something like the following:

Q is True	R is True
	R is False
Q is False	R is True
	R is False

It may even be helpful to think of it in these terms:



This way, we're getting all the possible combinations of Q:true, Q:false, R:true, and R:false. If this little explanation is confusing for you, it's probably best to move on. Perhaps it will make more sense later, and even if it doesn't, it's okay since this is sort of conceptual background work rather than something that is vital to understanding the truth table. What you need to understand at minimum is simply that by following this procedure, you're creating all of the possible combinations of T and F for the atomic sentence letters in the formulas you are trying to analyze using the truth table.

Now let's look at an eight row table. The first thing we do is start on the rightmost input (sentence letter) column, and alternate T and F every **one** row. Like so:

P	Q	R	**P***[Math Processing Error]***(Q***[Math Processing Error]***R)**	**R***[Math Processing Error][Math Processing Error]***Q**
Q R		T		
Q R		F		
Q R		T		
Q R		F		
Q R		T		
Q R		F		
Q R		T		
Q R		F		

Then we move one column to the left and alternate every **two.** Like this:

P	Q	R	**P***[Math Processing Error]***(Q***[Math Processing Error]***R)**	**R***[Math Processing Error][Math Processing Error]***Q**
Q R	T	T		
Q R	T	F		
Q R	F	T		
Q R	F	F		
Q R	T	T		
Q R	T	F		
Q R	F	T		
Q R	F	F		

Finally, we move to the left again and alternate every **four** rows so that we *double* the amount we are alternating by each time we move to the left. Like this:

P	Q	R	**P***[Math Processing Error]***(Q***[Math Processing Error]***R)**	**R***[Math Processing Error][Math Processing Error]***Q**
T **R**	T	T		
T **R**	T	F		
T **R**	F	T		
T **R**	F	F		
F **R**	T	T		
F **R**	T	F		

P	**Q**	**R**	**P**[Math Processing Error]**(Q**[Math Processing Error]**R)**	**R**[Math Processing Error][Math Processing Error]**Q**
F **R**	F	T		
F **R**	F	F		

The result is a truth table that's totally ready to solve: we have our columns labeled, our rows easily distinguishable, and most importantly we have all of the input side filled in in the standard way. This input side will be basically the same for every truth table. That is, you always follow this pattern:

Start below the rightmost atomic sentence letter and alternate every one row. Then move to the left one column and alternate going down every two rows. Finally, move to the left one last column and alternate going down that column every four rows. Extrapolate for bigger tables.

Solving a Truth Table

There are two methods to solving a truth table—filling in the right side or output side of the truth table. On the **Brute Force** method, you simply calculate each cell in the table by plugging the truth values of the sentence letters and working your way from the inside out. On the **Intuitive** method, you use your intuitive understanding of the operators to save yourself some work.

Brute Force Method

The Brute Force method to solving a truth table is simply to plug in the truth value for each individual letter by carrying them over from the input side and then calculating the truth value of each complex proposition given the truth values of the input side. This method is safer, so if you feel lost or lose confidence, then just revert back to the brute force method. It does have two downsides, though: a) It requires a lot more work and so takes more time, and b) it involves more individual steps and so the probably of making a simple mistake increases a bit. The second problem is probably balanced out by the riskiness of the Intuitive Method.

Write the truth value of each ***letter*** as it appears on the left side of the table under each letter as it appears on the right side

1. Starting with the "inner most" operators (inside the most parentheses) calculate the truth value of the whole relationship.

2. Then work your way out until you've calculated the truth value of the main operator.

So, step 0 is to make the truth table, as discussed in Basic Symbolization:

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
T	T		
T	F		
F	T		
F	F		

Step 1 is to fill in the truth values on the right side for the sentence letters

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
F	T	T T T	T T T
F	F		
T	T		
T	F		

Step 2 is to find the inner most operators (the ones with the most amount of parentheses outside of them.

P	Q	(P *[Math Processing Error] [Math Processing Error]*)	(Q *[Math Processing Error] [Math Processing Error][Math Processing Error]*)
F	T	T T T	T T T
F	F		
T	T		
T	F		

And then solve for those values using the truth tables that we use to define each operator.

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error] [Math Processing Error]*(P*[Math Processing Error]*Q))
T	T	T T	T T
T	F		
F	T		
F	F		

Repeat the steps in 2 (working from "in" to "out") until you've solved the truth value of the main operators (the operators that are only inside the outermost parentheses. Keep in mind that if there are no outermost parentheses, then they are implied).

So Q *[Math Processing Error]* (P*[Math Processing Error]*Q) is actually supposed to read: (Q *[Math Processing Error]* (P*[Math Processing Error]*Q))

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error] [Math Processing Error]*(P*[Math Processing Error]*Q))
T	T	T T	T F
T	F		
F	T		
F	F		

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error] [Math Processing Error]*(P*[Math Processing Error]*Q))
T	T	T	F
T	F		
F	T		
F	F		

Repeat in each row:

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error] [Math Processing Error]*(P*[Math Processing Error]*Q))
T	T	*[Math Processing Error]* T *[Math Processing Error] [Math Processing Error]*	*[Math Processing Error]* F *[Math Processing Error] [Math Processing Error]*
T	F	T F T	F T F
F	T	F T F	T F T

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
F P	F Q	(P [Math Processing Error] F F F (Q[Math Processing Error]P))	(Q [Math Processing Error] F F F [Math Processing Error](P[Math Processing Error]Q))
T P	T Q	(P [Math Processing Error] T (Q[Math Processing Error]P)) [Math Processing Error] [Math Processing Error]	(Q [Math Processing Error] F [Math Processing Error](P[Math Processing Error]Q)) [Math Processing Error] [Math Processing Error]
T T	T F	[Math Processing Error] T F T [Math Processing Error]	[Math Processing Error] F F F [Math Processing Error]
T F	T F	F T T F F	T F F T T
F F	T F	F F F F F	T T F F F
F	F	F T	F F

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
T	T	[Math Processing Error] T [Math Processing Error]	[Math Processing Error] F [Math Processing Error]
T	F	T F	F F
F	T	F F	T F
F	F	F T	F T

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
T	T	T	F
T	F	F	T
F	T	T	F
F	F	T	F

Then you analyze the results by looking at the right side of the truth table!

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))
T	T	**T**	**F**
T	F	**F**	**T**
F	T	**T**	**F**
F	F	**T**	**F**

Intuitive Method

The Intuitive Method to solving a truth table uses our intuitive understanding of the logical operators and their individual truth tables to save as much work as possible. We can often eliminate half of our work for one column in a single swoop.

It is a faster method, but also increases the chances that we'll make a mistake by moving too quickly and missing something or being overconfident and ignoring important details.

How does it work? Let's start with an example and then we'll look at how it works a bit more.

P	Q	(P [Math Processing Error] (Q[Math Processing Error]P))	(Q [Math Processing Error] [Math Processing Error](P[Math Processing Error]Q))	(Q [Math Processing Error] [Math Processing Error]P)

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T			
T	F			
F	T			
F	F			

Starting with the first column, we ask ourselves what we know about the main operator (*[Math Processing Error]*). I notice that there is only one atomic letter as the antecedent to this conditional. I know that if the antecedent to a conditional is false the whole conditional is true (since it's only false when T*[Math Processing Error]*F and if it's F*[Math Processing Error]*?, then it's certainly not T*[Math Processing Error]*F!). So I just need to find the rows on which P is false and I know the whole first formula will be True!

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T			
T	F			
F	T	T		
F	F	T		

Then we solve the rest more-or-less using the Brute Force method. When is (Q*[Math Processing Error]*P) true? When they're the same! It's false if they're different.

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T	T		
T	F	F		
F	T	T		
F	F	T		

Now we look at the second column. When is a conjunction true? It's only true in one case: when both conjuncts are true. So if Q is false...? Then the whole second formula (Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q)) is false.

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T	T		
T	F	F	F	
F	T	T		
F	F	T	F	

Now that we know Q is true in the remaining cells of the second output column, we need to ask ourselves when *[Math Processing Error]*(P*[Math Processing Error]*Q) is true. It says "neither P nor Q are true". When would that be true? Only when P and Q are both false (remember it would be equivalent to (*[Math Processing Error]*P*[Math Processing Error][Math Processing Error]*Q)). Are both false in rows 1 or 3? Nopers. So that means that *[Math Processing Error]*(P*[Math Processing Error]*Q) is false in both of our remaining rows. If just one conjunct is false, the whole conjunction is false. So:

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T	T	F	
T	F	F	F	
F	T	T	F	
F	F	T	F	

Okay, final column. Another way of doing the intuitive method is simply to understand what the formula says in a more intuitive way than the logical formula. (Q *[Math Processing Error]* *[Math Processing Error]*P) says something like "Q is true and P is false." When we understand it this way, it's easier to figure out on which row(s) it will be true: we're looking for the row(s) where Q is true and P is false. Which row is that?

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T	T	F	
T	F	F	F	
F	T	T	F	T
F	F	T	F	

Now we just fill in the rest as false since we know that on those rows P isn't true while Q is false.

P	Q	(P *[Math Processing Error]* (Q*[Math Processing Error]*P))	(Q *[Math Processing Error]* *[Math Processing Error]*(P*[Math Processing Error]*Q))	(Q *[Math Processing Error]* *[Math Processing Error]*P)
T	T	T	F	F
T	F	F	F	F
F	T	T	F	T
F	F	T	F	F

Okay, now that we've gone through the intuitive method, let's take a look at some of the rules which we can use while doing the intuitive method. Here are the rules I use:

- Antecedent **F** or Consequent **T** *[Math Processing Error]* Whole Implication **T**
- One disjunct **T** *[Math Processing Error]* Whole Disjunction **T**
- One conjunct **F** *[Math Processing Error]* Whole Conjunction **F**
- "*[Math Processing Error]*Q" means "Q is **false**".
    - So "(*[Math Processing Error]*P*[Math Processing Error]*Q)" means "P is **false** and Q is **true**"

These four rules can save you loads of time. Just identify which is simplest in a formula (surrounding the *main operator*): an antecedent? A consequent? A disjunct? A conjunct? Then you identify when that element fits the intuitive rule and eliminate lots of work!

Analyzing a Truth Table

Classification

If you're being asked to **analyze a single proposition** using a truth table, then you automatically know that the answer will be one of three options. This is called a Classification problem because you're classifying a single proposition.

1) **Tautology/Tautologous**: the column under the proposition is filled only with Ts.

A "Logical Truth" or Tautology is a statement that, regardless of how the world turns out to be, will be true. Think of "we'll either have a democratic president or we won't have a democratic president." Even if the world ends and we have no president at all, that disjunctive statement is still true!

Example [Math Processing Error]

T
T
T
T

2) **Self-Contradiction/Self-Contradictory**: the column under the proposition is filled only with F's.

A self-contradiction is always false no matter how the world ends up being. Think of the example "Kamala Harris is going to be our next president, but luckily we won't have to have Kamala Harris as our next president." It doesn't matter what actually happens in the world, this statement will always be false. It can't possibly be true because it contradicts itself.

Example [Math Processing Error]

F
F
F
F

3) **Contingent**: the column under the proposition is filled with a mixture of "T"s and "F"s.

Contingent propositions are true or false depending on how the world is. The simplest contingent propositions are atomic propositions, which either describe the world accurately or don't describe the world accurately.

Example [Math Processing Error]

F
F
T
F

Comparison

If you're being asked to **analyze a set of propositions** using a truth table, then you know that the answer will be one of the following four options. This is called a Comparison problem because you're comparing multiple propositions with one another in order to determine what logical relation holds between them. This is the most complex type of truth table analysis problem.

1) **Logically Equivalent**: each row of the *output* side of the truth table is the same on each column. So each row is a homogeneous set of either all T's or all F's. Logically equivalent propositions are true and false in exactly the same states of the world: they give the same output every time.

Example [Math Processing Error]

T	T
F	F
T	T
F	F

2) **Contradictory**: the truth values are the *opposite* on each row of the output side of the truth table. Notice that, since we only have two truth values (T and F), that means that we could only ever have contradictory *pairs* of statements. A set of three or more couldn't possibly be contradictory.

Example [Math Processing Error]

T	F
F	T
F	T
T	F

3) **Consistent**: Once you've determined that a set of statements is neither contradictory nor logically equivalent, you should check to see whether it is consistent. A consistent set of propositions is one where the logical form or structure of those propositions allows them to all be true at the same time. If the world is a certain way, then all of the statements will end up being true. So when testing for consistency, you're simply looking at the output side of the truth table for a row completely filled with T's. If you find it, then that set is consistent: they can all be true at the same time.

Example [Math Processing Error]

T	T
F	T
F	F
T	F

4) **Inconsistent**: The last option is to call a set of propositions inconsistent. If you never find that row filled completely with T's, then the propositions you are analyzing are inconsistent: they cannot, as a matter of logical structure, be true at the same time, no matter the state of the world.

Example [Math Processing Error]

T	F
F	F
F	F
T	F

Note that a set of logically equivalent propositions is likely to be consistent since it's likely to have one line on which all propositions are true. Furthermore, *every* contradictory pair of propositions is inconsistent. For the sake of the class, instructors will often require that you choose *only one* of the four options on a multiple-choice quiz or exam, though, so how do you decide?

Easy: just test for these four relations in order. Start by asking "are they logically equivalent?" Then, when you've found a line on which they have different truth values, as yourself "could they be contradictory?" Next, if you decide that they aren't contradictory (or couldn't be because it's a set of 3 or more), go hunting for one row on which all of the columns have a T. If you find it, then the set is consistent. If you never find such a row, then the set is inconsistent.

> **In short**: always answer with the **strongest** answer available. If a set of statements is *both* logically equivalent *and* consistent, then the correct answer on a multiple-choice test is "Logically Equivalent" since that's a stronger claim (it's a claim about *every* row rather than just one row).

Testing for Validity

If you're being asked to **analyze an argument or inference**, then you know that there are only two possible answers: Valid and Invalid. It's easier to start by discussing an invalid inference:

1) **Invalid**: an inference or argument is invalid if you find a row of the output side of the truth table where all of the premises are true, but the conclusion is false. So this, depending on how many propositions make up the argument, will typically be a row that looks like TTTTF, TTF, TF, TTTF, etc. If you find even just one row where the output side looks like this, then you've proven that the argument is invalid. *I like to think of it as searching for a radioactive row. If you find one with *all* true premises and a false conclusion, then you've found a radioactive row and therefore you've found out that the argument is invalid.*

Figure *[Math Processing Error]*: A Geiger counter for detecting atomic radiation

> I like to think of the counterexample line—the line that tells you the inference is invalid— as a sort of "radioactive" line you're searching for. Think of it like this: you've got your Geiger counter and you're scanning through the truth table for radiation. If you find a radioactive line, then the argument is bad (invalid). If you don't find a radioactive line, then the argument is clean (valid).

2) **Valid**: If you look through the whole output side of a truth table and never see a row where *every* premise is true and the conclusion is false, then you've found a valid argument. Remember that rows where one premise is false don't count and rows where it's all false premises and a true conclusion don't count. Everything is safe except rows somewhat similar to TTF or TTTTF or the like, depending on how many premises and conclusions there are.

Validity means it's impossible that the premises would be true while the conclusion is false. So if you find a row (even just one!) in a truth table telling you that if the world is like this (what you see on the input side of the row) then the premises will be true while the conclusion is false. This would never happen with a valid argument, so it follows that the argument you've found is an invalid argument.

The Reverse Truth Table Method

What happens if we symbolize or translate an argument and we end up with 5, 6, or more atomic sentence letters? Are we doomed to create a truth table with 32, 64, or more rows? That would be a fate worse than many things!

Fear not, dear student. We have a method for directly testing the validity of an inference without having to build a complete truth table. It's called the **indirect** or **reverse truth table method.** The basic idea is to assign "invalid" truth values to the premises and conclusions, figure out what we need the atomic sentence letters to be in order for those truth values to obtain, and then see if we can consistently apply truth values to the atomic sentence letters in order to create a counterexample (a line with all true premises and a false conclusion). So, basically, we're looking for that "radioactive" row (TF, TTF, TTTF, etc.) from the complete truth table from the previous section, but instead of building a whole truth table, we're simply going right to the end and testing if there *could be* such a row. We're testing to see if a radioactive row is even possible.

So, let's walk through an example, and then we'll come up with a set of steps for doing the reverse truth table method. Here's the English argument:

If you don't pass the driver's license written test, then you won't have a driver's license (until you are able to pass it). But I don't have a driver's license, so that means I didn't pass the driver's license written test?

Consider how confusing this would be to someone who has passed the written portion of the test, but hasn't yet completed the driving test. They did pass the written test, but still don't have a driver's license! This is confusing, because this argument is what is called "Affirming the Consequent". It's a formal fallacy, or an invalid argument that might appear to be valid. Let's symbolize it (ignoring the parenthetical):

~W [Math Processing Error] ~L

~L

[Math Processing Error] ~W

> Remember that "*[Math Processing Error]*" means "therefore"

In order to run the reverse truth table method of analysis on this argument, we'll first want to set up as if we are doing the *output* side of a truth table. We'll want to give ourselves lots of room to work with:

(~W *[Math Processing Error]* ~L) / ~L // ~W

Now, the next thing to do is to assign truth values to each **whole proposition** so that we have a radioactive row or counterexample row. In this case, there are two premises and a conclusion, so the counterexample row will be TTF:

(~W *[Math Processing Error]* ~L) / ~*[Math Processing Error]* // *[Math Processing Error]*

The setup part is done. Now we need to do the hard work: actually work out if we can consistently apply truth values to the atomic sentence letters. We'll go one step at a time, starting with the conclusion. Why start with the conclusion? It's typically easier to make a sentence false given the truth tables for disjunction, negation, and conditional; so the conclusion generally is the easier place to start. Can you figure out why the rules for these operators make establishing falsehood easier?

(~W *[Math Processing Error]* ~L) / ~*[Math Processing Error]* // *[Math Processing Error]*

The conclusion is false, so W will need to be true since the conclusion is ~W. So then we assign T to every W that appears in the argument:

(~*[Math Processing Error] [Math Processing Error]* ~L) / ~*[Math Processing Error]* // *[Math Processing Error]*

And then work out how that affects the formulas containing the atomic letters I just assigned truth values to. In this case, if W is true, then ~W is false, and if an antecedent is false, then the whole conditional is true. The main operator is the conditional, and so no problems with the first premise:

(*[Math Processing Error]* ~L) / ~*[Math Processing Error]* // *[Math Processing Error]*

What about the second premise? Well, since we haven't yet been forced to assign a truth value to L, we can assign whatever we want to it, so we'll make it false! The result will be that ~L is true.

(*[Math Processing Error]* ~L) / *[Math Processing Error]* // *[Math Processing Error]*

At this point, we've symbolized the argument, put it in a row as if we were going to make a truth table out of it, and then tried assigning truth values to the atomic sentence letters to make a radioactive row (in this case TTF). Since we were able to do so without coming across a contradiction, we know that the inference is **invalid**.

Here's a sort of algorithm for the reverse truth table method of analysis:

1. Symbolize the inference and write out in a single row using slashes between formulas.

2. Assign truth values to each complete formula such that **all premises are true** and **the conclusion is false.**

3. Then **calculate the truth value of each atomic sentence letter** given the truth value of the whole formula. Start with the most restrictive formulas. If you run into a case where multiple truth values would work, then start a new line for each possibility and test each going forward.

4. Then transfer the atomic sentence letter truth value(s) to the other instances throughout the whole inference (transfer the truth values of, for example, 'A' to all other A's throughout the formula).

5. Then calculate whether it is possible to continue assigning truth values to atomic sentence letters and transferring those values to other instances **without running into a contradiction** (that is, where a single letter must be both T and F)**.**

6. If you run into **no contradiction**, then the inference is invalid *because the radioactive row is possible*. The process is over.

If you run into **a contradiction**, then possibly the inference is valid. Complete all lines you've started to ensure that there is no possible consistent assignment of truth values. You only need one possible row where there is no contradiction to prove that the inference is invalid; whereas you need to eliminate every possible row that could have true premises and a false conclusion before you can know it's valid.

Okay, let's try a more complex one:

Lila: We're either going home or I'm going home alone unless you both assure me you will drive us home later and will phone the babysitter.

Diego: I can't assure you that I'll be sober enough to drive us home later.

Lila: Well, I'm not going home alone and you're not staying the night here.

Diego: Well, then, I guess either we're both going home now or we're getting a motel room.

If we conceive of the whole exchange as one big inference, we can **symbolize** it the following way:

(H *[Math Processing Error]* (I *[Math Processing Error]* (A *[Math Processing Error]* P))) / ~A / (~I *[Math Processing Error]* ~S) // (H *[Math Processing Error]* M)

Now, we **assign truth values to the premises and conclusion** so that they'll be radioactive.

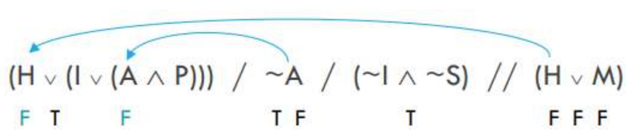
Here, I've gone ahead and put them under the main operators:

(H *[Math Processing Error]* (I *[Math Processing Error]* (A *[Math Processing Error]* P))) / *[Math Processing Error]*A / (~I *[Math Processing Error]* ~S) // (H *[Math Processing Error]* M)

Next, I'm going to start looking for some simpler formulas **to assign truth values** to. I'm eyeing the conclusion and the ~A. The conclusion will only admit of one set of truth values: both H and M must be false for the *[Math Processing Error]* to be false. A must be false for ~A to be true.

(H *[Math Processing Error]* (I *[Math Processing Error]* (A *[Math Processing Error]* P))) / *[Math Processing Error]* / (~I *[Math Processing Error]* ~S) // (*[Math Processing Error]*)

Next, you **transfer truth values** from the ones you just assigned to all identical letters.

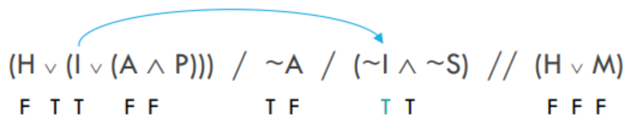


Notice how I'd need the right disjunct (I *[Math Processing Error]* (A *[Math Processing Error]* P)) to be true for the first premise to turn out true.

But we already know that A is false from premise 2. So that means the conjunction (A *[Math Processing Error]* P) must be false. Now that we know this, we must conclude that I is true for the first premise to turn out true. So we can conclude that I is true.

(*[Math Processing Error]* P))) / *[Math Processing Error]* / (~I *[Math Processing Error]* ~S) // (*[Math Processing Error]*)

Now we transfer the new atomic truth value:



Now let’s try to work out that third premise. First we can process the negation on I:

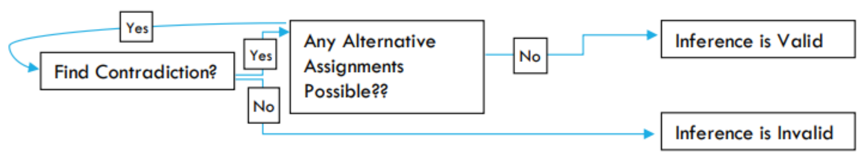
([Math Processing Error] P))) / [Math Processing Error] / ([Math Processing Error] ~S) // ([Math Processing Error])

We’ve already got a contradiction!!! Ouch! ~I would have had to be true for (~I [Math Processing Error] ~S) to turn out true. Both conjuncts need to be true. But ~I is false according to our assignment of false to I. Bummer dude!

What now? We’ve reached a contradiction! At this point we ask ourselves: “was there any step I made that I wasn’t *forced* to make?” In this case, no, we didn’t arbitrarily choose true or false for any letter, so every step we took was forced by logic. That means there aren’t any alternative assignments to consider and therefore **the radioactive counterexample is impossible.** Our inference is **valid**.

You could have also assigned values to I and S given the third premise, but I chose to finishing the first premise. Either way would’ve resulted in a contradiction. *This is not an example of an alternative assignment*. We’ll cover alternative assignments below:

Here's a handy flow chart for you:



Let’s try a quick example with alternative assignments possible:

(H [Math Processing Error] (I [Math Processing Error] (A [Math Processing Error] P))) / [Math Processing Error]A / (H [Math Processing Error] M) // (~I [Math Processing Error] ~S)

There are many ways for Premise 1 to be true, many ways for Premise 3 to be true, and many ways for the Conclusion to be false (an exception to the generalization I made earlier). So we’re not being forced as much as we were in the previous example. Arguments where the conclusion has more than one way of being false are typically the hardest arguments to do using the reverse truth table method. Hardest, that is, in terms of how much work is involved. Remember that this isn’t difficult in the sense of it being a complicated procedure. It’s not like playing chess. We could program a computer to do this whole method in an afternoon. It does, though, sometimes take a bit of work to work out the answer. Let’s start with what we are forced to do:

(H [Math Processing Error] (I [Math Processing Error] ([Math Processing Error] [Math Processing Error] P))) / [Math Processing Error] / (H [Math Processing Error] M) // (~I [Math Processing Error] ~S)

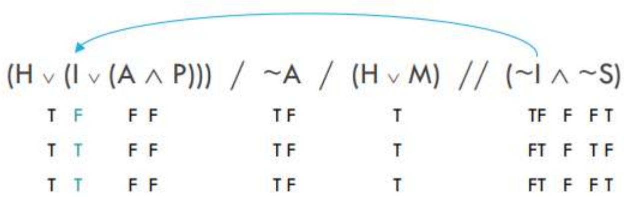
A must be false because ~A is true. At this point, we aren’t forced to do anything for Premise 1 since there are still many ways for it to come out true. Nothing has changed for Premise 3 and the Conclusion. At this point we need to split our line into all possible successful assignments. I’m going to start with the conclusion (I chose this more or less arbitrarily). Here’s what I do:

[Math Processing Error]

If you focus on the conclusion, it looks sort of like a truth table now, doesn’t it? Now we complete the process for each possible line:

[Math Processing Error]

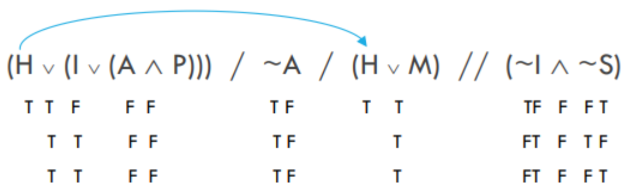
Now I’m transferring truth values:



And then calculate what needs to happen given the changes you've made. I've changed the first premise, so now I notice that (A *[Math Processing Error]* P) is false and now I is false in my first row, so that means H must be true. In the other rows, nothing has yet forced me to assign anything to H.

[Math Processing Error]

Then transfer that H:



At this point, nothing is forcing my hand. If you look carefully, you'll see that there is no single letter on any single line that *must* be a given truth value for the truth values we've assigned the complete formulas to obtain. So now what???

If we are free to assign any truth values we want, then we aren't going to run into any contradictions. It follows that there are no contradictions in these three rows and there is therefore *at least one row* where there are no contradictions. This inference is **invalid**.

> All we need is **one** row where there are no contradictions to prove that the inference is invalid.

That's all! That's how we do the reverse truth table method. One can imagine how to use this to test even super complex sets of sentences for consistency: just assign "true" to each individual sentence and then look for a contradiction. One line with no contradiction? You've got a consistent set of sentences.

7.6: Conclusion

We'll move onto Natural Deduction in the next chapter, and that's a continuation of propositional logic. I'll have a lot more to say about logic in general, its usefulness, and its value at the end of Chapter 7. If you won't be studying natural deduction, please do go and read sections 7.5 and 7.6 as I think they are important things to think about after learning a bit of deductive logic. For now, let me just congratulate you! You've learned how to take English sentences, symbolize them into formulas in propositional logic as a means of symbolizing the logical form of those sentences, and then you learned how to use those symbols to learn things about the original sentences you were symbolizing. You can now test for the validity of inferences and the consistency of sets of propositions, among many other things. I mention these two because I think they're probably the most useful of the set of analyses you can now run on an argument or set of statements. That's all for now!

7.E: Chapter Six (Exercises)

Exercise [Math Processing Error]: Basic Symbolization

Symbolize the following ordinary language sentences into propositional logic using the five basic logic symbols, sentence letters, and parentheses/brackets.

A. I used to go to the gym every day, but now my kids keep me busy.

B. I used to go to the gym every day, but now if my kids are awake, I'm usually with them.

C. I am the manager of this team, and what I say goes.

D. I'm not the manager of this team, but my ideas should be taken seriously.

E. I'm not going to marry you if you won't even clean up your apartment when I come over.

F. We need to get a divorce, or at the very least something needs to change.

G. If I'm going to stay with you, then you need to treat me with respect and honor my wishes.

H. No people here are warm enough.

I. Either you light a fire, or I will light one instead

J. Either you light a fire, or if you don't, I will have to do it.

K. I hate being on a diet whenever, but luckily only whenever it's a calorie-restriction diet. (think "if, but luckily only if").

Exercise [Math Processing Error]: More Complex Symbolization

Symbolize the following into propositional logic. Be careful with different forms of "not both", "neither nor", "not x and not y", "either not x or not y", as well as different conditional forms like "A if B", "If A then B", "A only if B", "only if A, will b", and so on.

A. I'm either not going to go out with you tonight, or not going hiking with you tomorrow.

B. I go to the gym every day, but if I become a parent, I won't go to the gym everyday.

C. If and only if you clean your room will you get your allowance and be allowed to go out with friends later.

D. I want you to come out with us, but I don't want you to both feel uncomfortable and make others uncomfortable.

E. Peace will finally reign only if those powerful folks who would make soldiers into expendable pawns and those who would make enemies into monsters are stripped of their power.

F. If a judge is to rule fairly on constitutional matters, then neither may they have conflicts of interest, nor may they allow their training in constitutional law to lapse.

G. Neither will we buy a new house if there isn't one available nor will we move to a new apartment if the rent is too high.

H. Either he leaves, or I do unless he apologizes and offers financial restitution.

Exercise [Math Processing Error]: Framing Words

Use what you've learned about the framing words we use in English to mark off logical structures like "either...or...", "if...then...", and "both...and..." to symbolize the following complex English sentences:

A. If you don't learn grammar early on in life, then it becomes harder to both recognize grammatical structures and either succeed academically or at least reason logically.

B. Terrible is the day that we surrender if we haven't both given the battle our all and been just in our dealings with our opponents.

C. If and only if we are both fair in our negotiations and we neither give the appearance of disrespect nor betray the fact that we are in fact loyal to the crown, will we either be successful in our negotiations or at least not be punished for failing to negotiate successfully.

D. Provided we will not win this trade war, we will lose if and only if we either fail to compromise and fail to achieve some semblance of a draw, or we both give up too much in our negotiations and walk away with less than we had at the beginning.

E. Only if we are able to win over the citizens' hearts and win over their minds, will our regime-change operations be successful and will our sacrifices not have been in vain.

Exercise *[Math Processing Error]*: Main Operators

Identify the Main Operator of each propositional logic formula. Many of these are from Carnap.io/book CC-BY 4.0 Intl License.

A. (P *[Math Processing Error]* Q)

B. ((P *[Math Processing Error]* R) *[Math Processing Error]* (P *[Math Processing Error]* (Q *[Math Processing Error]* *[Math Processing Error]*S)))

C. ((Q *[Math Processing Error]* (R *[Math Processing Error]* S)) *[Math Processing Error]* ((P *[Math Processing Error]* Q) *[Math Processing Error]* (P *[Math Processing Error]* R)))

D. (((R *[Math Processing Error]* S) *[Math Processing Error]* (P *[Math Processing Error]* *[Math Processing Error]*Q)) *[Math Processing Error]* (*[Math Processing Error]*R *[Math Processing Error]* S))

E. (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* P))))

F. [*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* P]

G. *[Math Processing Error]*(P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* P))))

H. (*[Math Processing Error]*((R *[Math Processing Error]* S) *[Math Processing Error]* *[Math Processing Error]*(P *[Math Processing Error]* *[Math Processing Error]*Q)) *[Math Processing Error]* *[Math Processing Error]*(*[Math Processing Error]*R *[Math Processing Error]* S))

I. (Q *[Math Processing Error]* (((R *[Math Processing Error]* S) *[Math Processing Error]* (P *[Math Processing Error]* Q)) *[Math Processing Error]* (P *[Math Processing Error]* R)))

(Careful: parentheses are different from problem C)

Exercise *[Math Processing Error]*: Computing Truth Values

Compute the truth values of the following formulas, where A-L are all true and M-Z are all false.

A. (A *[Math Processing Error]* Q)

B. ((P *[Math Processing Error]* B) *[Math Processing Error]* (L *[Math Processing Error]* (Q *[Math Processing Error]* *[Math Processing Error]*S)))

C. ((Q *[Math Processing Error]* (R *[Math Processing Error]* F)) *[Math Processing Error]* ((D *[Math Processing Error]* Q) *[Math Processing Error]* (C *[Math Processing Error]* R)))

D. (((R *[Math Processing Error]* S) *[Math Processing Error]* (P *[Math Processing Error]* *[Math Processing Error]*Q)) *[Math Processing Error]* (*[Math Processing Error]*R *[Math Processing Error]* S))

E. (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* P))))

F. (P *[Math Processing Error]* (B *[Math Processing Error]* (C *[Math Processing Error]* (D *[Math Processing Error]* P))))

G. [*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* P]

H. [*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* K]

I. *[Math Processing Error]*(P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* P))))

J. (P *[Math Processing Error]* *[Math Processing Error]*(B *[Math Processing Error]* (*[Math Processing Error]*C *[Math Processing Error]* *[Math Processing Error]*(D *[Math Processing Error]* P))))

K. (*[Math Processing Error]*((R *[Math Processing Error]* S) *[Math Processing Error]* *[Math Processing Error]*(P *[Math Processing Error]* *[Math Processing Error]*Q)) *[Math Processing Error]* *[Math Processing Error]*(*[Math Processing Error]*R *[Math Processing Error]* S))

L. (*[Math Processing Error]*((D *[Math Processing Error]* S) *[Math Processing Error]* *[Math Processing Error]*(E *[Math Processing Error]* *[Math Processing Error]*Q)) *[Math Processing Error]* *[Math Processing Error]*(*[Math Processing Error]*B *[Math Processing Error]* S))

M. ((Q *[Math Processing Error]* ((R *[Math Processing Error]* S)) *[Math Processing Error]* (P *[Math Processing Error]* Q) *[Math Processing Error]* (P *[Math Processing Error]* R)))

N. ((B *[Math Processing Error]* (R *[Math Processing Error]* E)) *[Math Processing Error]* ((A *[Math Processing Error]* Q) *[Math Processing Error]* (P *[Math Processing Error]* R)))

Exercise *[Math Processing Error]*: Truth Tables Classify

Build a Truth Table to Classify the following single propositions:

A. *[Math Processing Error]*(P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* (P *[Math Processing Error]* P))))

B. (((R *[Math Processing Error]* S) *[Math Processing Error]* (R *[Math Processing Error]* *[Math Processing Error]*S)) *[Math Processing Error]* (*[Math Processing Error]*R *[Math Processing Error]* S))

C. ((P *[Math Processing Error]* R) *[Math Processing Error]* (P *[Math Processing Error]* (R *[Math Processing Error]* *[Math Processing Error]*P)))

D. [*[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(E*[Math Processing Error]* (F *[Math Processing Error]* D))] *[Math Processing Error]* E]

E. (((P *[Math Processing Error]* Q) *[Math Processing Error]* R) *[Math Processing Error]* ((P *[Math Processing Error]* R) *[Math Processing Error]* (Q *[Math Processing Error]* R)))

Exercise *[Math Processing Error]*: Truth Tables Compare

Build a truth table to compare the following sets of propositions. Many of these are from Carnap.io/book CC-BY 4.0 Intl License.

A. [~(B*[Math Processing Error]*E)*[Math Processing Error]*I] / [(B*[Math Processing Error]*E)*[Math Processing Error]*~I]

B. [*[Math Processing Error]*[Z *[Math Processing Error]* *[Math Processing Error]*(X*[Math Processing Error]* (Z *[Math Processing Error]* Q))] *[Math Processing Error]* Q] / Z / (X *[Math Processing Error]* Q)

C. (P *[Math Processing Error]* Q) / (~P *[Math Processing Error]* Q) / (~Q *[Math Processing Error]* ~P)

D. ((P *[Math Processing Error]* R) / (P *[Math Processing Error]* (R *[Math Processing Error]* *[Math Processing Error]*P)))

E. *[Math Processing Error]*[D *[Math Processing Error]* *[Math Processing Error]*(E*[Math Processing Error]* (E *[Math Processing Error]* D))] / [*[Math Processing Error]*D *[Math Processing Error]* (*[Math Processing Error]*E*[Math Processing Error]* (E *[Math Processing Error]* D))]

F. (P *[Math Processing Error]* Q) / (Q *[Math Processing Error]* R) / (P *[Math Processing Error]* R)

? Exercise *[Math Processing Error]*: Truth Tables Validity

Build a truth table to test each of the following inferences for validity. Many of these are from Carnap.io/book CC-BY 4.0 Intl License.

A. (P *[Math Processing Error]* R) // (*[Math Processing Error]*P *[Math Processing Error]* *[Math Processing Error]**[Math Processing Error]*R)

B. P // (P *[Math Processing Error]* *[Math Processing Error]*(P *[Math Processing Error]* P))

C. (P *[Math Processing Error]* Q) // (Q *[Math Processing Error]* P)

D. A / (A *[Math Processing Error]* B) // (B *[Math Processing Error]* A)

E. D / (C *[Math Processing Error]* (C *[Math Processing Error]* (C*[Math Processing Error]* D)))

F. *[Math Processing Error]*P / (P *[Math Processing Error]* R) // (R *[Math Processing Error]* *[Math Processing Error]*(P *[Math Processing Error]* P))

G. ((P *[Math Processing Error]* Q) *[Math Processing Error]* (P *[Math Processing Error]* R)) // ((P *[Math Processing Error]* Q) *[Math Processing Error]* *[Math Processing Error]*(P *[Math Processing Error]* R))

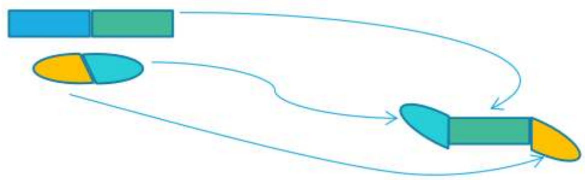
CHAPTER OVERVIEW

8: Natural Deduction

8.1: What is it?

Natural deduction is a *proof system* for propositional logic. In short, it's a way of proving that really complex argument forms are **valid** by using really simple argument forms that are easy to understand. In other words, first we learn a series of "rules." Each rule is an inference that we know is valid. Then we use this method to discover that much more complex arguments are in fact made up of these little arguments we call "rules." **Complex arguments can be proven, in a step-by-step way, to be valid by deriving their conclusions from their premises** (by moving step-by-step from just their premises to their conclusions in logically valid steps). If we get from A to B using only valid steps, then the move directly from A to B is itself valid!

> *Natural deduction is, in practice, a sort of **game**. We are given some pieces at the beginning (premises), along with rules for how to split apart, combine, and transform those pieces, and we try different strategies until we win the game by making the conclusion. It's sort of like a puzzle, or a step-by-step strategy game.*



Let's dive right in with an example:

Example *[Math Processing Error]*

I don't believe in ghosts, but if I did believe in ghosts, that sound outside would be frightening. Luckily, that sound isn't frightening to me!

So I guess you really don't believe in ghosts!

The first step is to *symbolize* or *formalize* the argument. This is the process where we take out all of the logically irrelevant content of the argument as it is written here, leaving only the argument *structure* or *form* and some symbols which stand in for the content that we took away. This is the first thing we learned in Propositional logic in Chapter 6.

[Math Processing Error] B *[Math Processing Error]* (B *[Math Processing Error]* F)

*[Math Processing Error]*F

[Math Processing Error] *[Math Processing Error]* B

The little "*[Math Processing Error]*" symbol means "therefore" or "what came before is an argument for the following conclusion."

Now this argument is really clearly valid just by looking at it[1] but it's helpful to start with a really easy example so that we can build up to the really tricky examples. First, look at the "*[Math Processing Error]* B" in the first premise. Notice how it's identical to the conclusion? Well, anything implies itself deductively. So here's an argument that is very, very obviously valid:

[Math Processing Error] B

[Math Processing Error] *[Math Processing Error]* B

Obviously if "*[Math Processing Error]* B" is true, then "*[Math Processing Error]* B" *must* be true. As a matter of absolute necessity. That's what it means to be valid! Does that make sense?

Now let's look back at the original argument and try to figure out how it's valid. This time, I'll number the premises and put the conclusion off to the side to remind us what we're looking for. This is exactly the format of a standard natural deduction problem:

1. *[Math Processing Error]* B *[Math Processing Error]* (B *[Math Processing Error]* F)

2. *[Math Processing Error]*F / *[Math Processing Error]*B

The numbered propositions 1 and 2 are premises. The "[Math Processing Error]B" to the right of the slash is the conclusion of the argument. The goal, once we've got a problem in front of us, is to move from the premises to the conclusion *by means of a set of discrete steps*. The moves you can make are determined by which *rules* you know. We'll learn some rules soon and that will allow us to make a series of different moves from the premises forward towards the conclusion.

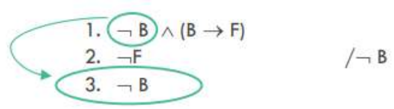
Notice how there are two B's in premise 1. So, we could take a complex route to proving the conclusion by isolating (getting it by itself) the "(B [Math Processing Error] F)" and then doing something else to get the B by itself and negating it—adding the "hook" or "[Math Processing Error]" in front of it. We have B and we want to get to not-B.

Here's a simpler way, for now:

If I tell you I have an apple and an orange, am I telling you that I have an apple? Yes. Am I telling you that I have an orange? Of course. If, alternatively, I tell you that "if I get into a fender bender on the way to work I'll be late, but if I don't, then I'll have to go to that awful meeting," am I telling you that "If I get into a fender bender on the way to work then I'll be late"? Yep! What about "If I don't get into a fender bender on the way to work, then I'll have to go to that awful meeting"? Yes again.

So as a general rule, it looks like any time I join two phrases with ***conjunction***, I'm allowed to *separate* them as well. Every time I tell you something involving 'but' or 'and,' I'm also telling you that everything on either side of the 'but' or 'and' is true independently from the rest of the sentence. We'll talk about this rule more later, but for now we'll call it ***simplification***, and move on. Let me reiterate, though, that when two things are joined by conjunction, we can just pull them apart in natural deduction and write them all by themselves.

Here's how we put that rule to use in our example above. I'll add a new line where I've derived a new statement from the existing premises. Our rule "simplification" lets me write something new on its own line. This is how all of the natural deduction rules work: they tell you to find something in the existing propositions and then once you find what you're looking for, you're allowed to write something new that the rule tells you to write. What I'm doing here is saying that given those premises (1&2) and all of the simple rules we know, I know it's logically okay or valid to posit or affirm or state the following (#3) all on its own:



And then I'll *justify* that move by appeal to a rule we just learned:

3. [Math Processing Error] B [Math Processing Error]

So what I'm communicating here is that I used premise 1 as my "raw material" and applying the rule of simplification to it, which allows me to isolate either the left conjunct (what's to the left of the wedge or "and" or conjunction symbol) or the right conjunct. I used the rule or operation "simplification" to get "[Math Processing Error] B" all by itself on a new line. What I've said so far is that given premise 1 and the rule "simplification", we know with deductive certainty that [Math Processing Error] B is true.

Okay, so that's one step in. Let's see how we're doing:

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F / [Math Processing Error] B

3. [Math Processing Error] B 1, simplification

Notice at this point that the line we've derive—line 3—is *exactly identical* to the conclusion.

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F / [Math Processing Error]

3. [Math Processing Error] 1, simplification

This is how we know we're done with a natural deduction derivation. Why? Well, simply because ***the whole goal of natural deduction is to derive the conclusion*** in a set of discrete, understandable steps.

In this case, it turned out to be a very simple derivation because the premises already basically say that B is false, and the conclusion is that B is false. Easy Peasy.

Moving on... We know lots of simple arguments are valid. Maybe we've used truth tables or some other simple proof method. Maybe we've instead just intuited that a simple argument *must* be valid like we did above with simplification. Either way, we've got a treasure trove of simple argument forms that we can use to prove much more complex argument forms to be valid.

How? Easy! By stringing together a series of simple steps—where each step is deductively valid—that get us from the premises to the conclusion, we will have thereby proven that *those* premises (or premises of that *form*) entail or imply with deductive certainty *that* conclusion (or conclusions of that *form*). If we can complete a proof like this, then we will have shown that if those premises are true, that conclusion *must* be true.

---

[1] That is, if you understand what the symbols means, which you should by now. If you don't, review the first and second Propositional Logic sections on translation and truth tables. It's important that you fully understand what each symbol means and can understand their inputs and outputs with ease before moving on to Natural Deduction.

8.2: Basic Rules of Implication

Okay, time to get a bit more complex. I'm going to introduce four basic rules so that we can do the derivation we just did in a more complex way. It turns out that those premises entail that conclusion in more ways than one!

Here's the first rule everyone learns. It's super straightforward. Don't overthink it:

> **Modus Ponens**
>
> *(Latin for "Mode of Affirming" or "Mode of Positing")*
>
> *If I find a conditional all by itself on one line, and then on another line I find the **antecedent** to that same conditional, then I am allowed to put the consequent on a new line and cite "MP" (modus ponens) as my justification.*
>
> *[Math Processing Error]*

I know what you're thinking. "But that's just what conditionals *are*!!!" They just mean that if the antecedent is true then so is the consequent. If you were thinking that. Great! That is in fact just what conditionals mean. Here, we just expand the meaning of a conditional out into an argument pattern.

Suppose someone reasoned as follows: "If he says "yes" to go to the dance with me, then I'll be so excited!" and then hours later, after asking this guy, says "He said "yes"! So now I'm so excited!" It would make perfect sense. There's a true conditional stating that if A happens, then B will happen, and then we confirm that A happened, so it follows that B will happen.

> *Modus Ponens allows us to break the consequent off of a conditional.*
>
> *There are many rules for manipulating conditional propositions,*
>
> *but this is the most basic: if you want the consequent, find the antecedent.*

How to read the rules of implication

Modus Ponens above looks like an argument, but is also how we write the first set of rules we learn: the **RULES OF IMPLICATION**. The rules of implication are all one-directional rules that are based on valid inference patterns—mostly inference patterns involving more than one premise. So we take these simple arguments that we know are valid and then turn them into rules that allow us to make moves in natural deduction.

The rules of implication have some set of premises above a line and a conclusion below the line. To read a rule of implication, you should interpret the formulas or patterns above the line as the **formulas you need to find** in the numbered propositions you have already derived. Then you look below the line for the corresponding **formula you should write on a new line**. Remember, every rule tells you something like "find x, write y." So if you find what's above the line, you then get to write what's below the line.

Furthermore, each rule is typically written using letters, but *those letters are stand-ins* for any formula at all. As a result, any of the following could be instances of Modus Ponens:

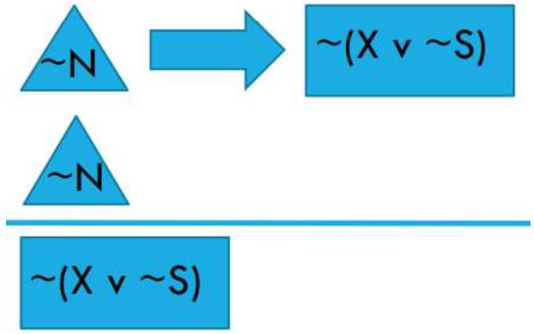
- *[Math Processing Error]*
- *[Math Processing Error]*
- *[Math Processing Error]*

What matters is that you can find one formula where the main operator is an arrow/horseshoe and then you can find another formula that is *identical* to the antecedent or left side of the arrow formula. If you find those, you can write down the right side or consequent of the arrow on a new line by itself.

It's helpful to some to think of the rules not in terms of variable letters, but instead in terms of shapes or abstract symbols that stand in for any well-formed formula at all. Here's how we might write Modus Ponens instead:

[Math Processing Error]

Perhaps it's easier to think of a triangle (delta) as a stand-in for any proposition at all. Then one can more easily recognize that these are all instances of this general structure or pattern:



If this is easier for you, then I encourage you to write the rules in your notes in symbols.

NEXT IS THE COMPLEMENT OF MODUS PONENS:

Modus Tollens

(Latin for "Mode of Denying")

If I find a conditional all by itself on one line, and then on another line I find the ***negated form of the consequent*** *to that same conditional, then I am allowed to put the negated form of the antecedent on a new line and cite "MT" (modus tollens) as my justification.*

[Math Processing Error]

Modus Tollens is a bit harder to understand, but it is still simply based on the definition of a conditional. We can look at the truth table for a conditional for starters:

P	Q	(P *[Math Processing Error]* Q)
T	T	T
T	F	F
F	T	T
F	F	T

Now, remember that the premises to the argument form of *modus tollens* above tell us that P implies Q and that Q is false. So we look at the rows in our table where Q is false, find that there's only one where "P implies Q" is true (row 4), and then look at what P is on that row. It is false. That's how we figure out *modus tollens* on a truth table.

We figure it out more intuitively by thinking that a conditional means "if the antecedent is true, then the consequent must be true." Another way of saying essentially the same thing is "if the consequent is false, then the antecedent must be false." Why? Well, if the consequent is false, then a true antecedent would give use "True *[Math Processing Error]* False." That would be bad, since we already know that "P *[Math Processing Error]* Q" is true and therefore cannot be "True *[Math Processing Error]* False" (P can't be true while Q is false).

Strategically, *modus tollens* allows us to **break the antecedent off of a conditional, but we have to negate it in the process**. So it's another rule for breaking apart a conditional: if you want the antecedent by itself, you'll need to find the negated consequent, and then you'll have to negate the antecedent when you bring it down to its own line. That's the price we pay for getting the antecedent on its own line.

Think about this: if Bruce Banner (The Hulk) says "If I get angry with you, then you will know it", then it follows that any time you can't tell that Banner is made at you is a time that he in fact is not mad at you. If he were mad at you, he'd be all green and scary looking. So if you can't tell that he's mad (if the consequent is false), then he certainly isn't mad at you (the antecedent is also false).

Or suppose it was true that "if you eat a bunch of food, then you'll be full." It follows that any time you aren't full, you won't have just eaten a bunch of food. Or "if you eat rotten food, you'll get a stomach ache." Suppose you don't have a stomach ache. Good news! You didn't eat rotten food.

The next rule deals with disjunctions instead of conditionals (we've had enough conditionals for one day).

Disjunctive Syllogism

(Latin for "Disjunctive Syllogism")[1]

If I find a disjunction all by itself on one line, and then on another line I find the ***negated form of either disjunct*** *to that same disjunction, then I am allowed to put the other disjunct on a new line and cite "DS" (disjunctive syllogism) as my justification.*

[Math Processing Error]

Now we might get super technical and say that the rule also involves a different argument form where the second premise is *[Math Processing Error]*P and the conclusion is Q. But we don't have to get too technical in this course. We can operate at a slightly more intuitive level and just recognize that we're allowed to take a disjunct out of a disjunction if we find the negation of the other disjunct.

Strategically, this rule allows us to **break apart a disjunction**, by pulling out one disjunct and putting it on a line of its own. If you want a disjunct, find the negation of the other disjunct. Simple, right?

The final rule of the first four we're introducing here is the following:

Hypothetical Syllogism

("Hypotheticals" and conditionals or implications are the same thing) If I find two conditionals—all by themselves on separate lines—where the antecedent of one is the consequent of the other, then I can "cut out the middle man" and write a new conditional on its own line which has the antecedent and conditional which aren't identical to one another. I justify this move by writing "HS" (hypothetical syllogism).

[Math Processing Error]

Again, this one makes some intuitive sense (hopefully). It's similar to the property of "transitivity" like you might have studied in a high school math class. In the form above, the Q is the "middle man" that we can take out if we find that it's useful.

Think in terms of the following formula: P *[Math Processing Error]* Q *[Math Processing Error]* R. If you imagine the arrow just going right through the "Q", you'll see the intuition behind hypothetical syllogism: the "middle man" is superfluous and can be removed.

This is a rule for **combining conditionals** by sticking the antecedent of one and the consequent of another together.

Note: the order of the premises doesn't matter, so the (Q *[Math Processing Error]* R) can show up first. That's not important. What *is* important is that the "middle man" appears on both the left and the right of the two premises. What this means is that the following is not an acceptable instance of hypothetical syllogism:

[Math Processing Error] ✗

Negative Buzzer Noise! That's not hypothetical syllogism. That's something else entirely, and it's ugly. More importantly, it's invalid and so won't work as a rule of inference.

Okay, here are those first four rules again:

4 Rules of Implication:

Modus Ponens	**Disjunctive Syllogism**
[Math Processing Error]	*[Math Processing Error]*

Modus Tollens	Hypothetical Syllogism
[Math Processing Error]	[Math Processing Error]

First Four Rules in Words:

Modus Ponens (MP): If you find the left side of an arrow, you get to write the right side

Modus Tollens (MT): If you negate the right side of an arrow, you get to negate the left side

Disjunctive Syllogism (DS): If you negate one side of a disjunction, you get to write the other disjunct.

Hypothetical Syllogism (HS): If you have two conditionals with a "middle man", you can take the middle man out.

Now let's go back to the beginning and play the original problem again, but this time we'll turn up the difficulty just a bit.

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F /[Math Processing Error] B

I cheated and already told you about simplification, which comes later on in our study of natural deduction, but we're going to use is again here.

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F /[Math Processing Error] B

3. (B [Math Processing Error] F) 1, simp

Okay, now where are we? We've got a "B [Math Processing Error] F" and we want "[Math Processing Error] B." Does that sound familiar? We want the negated form of the antecedent to a conditional. Look at the rules and try to figure out which rule we use next...

I'll wait...

Okay, you got it right. It's *modus tollens*! Here's how the completed derivation looks:

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F /[Math Processing Error] B

3. (B [Math Processing Error] F) 1, simp

4. [Math Processing Error] B 2, 3, MT

Again, I know I'm done when the last line I derived and the conclusion are **identical**:

1. [Math Processing Error] B [Math Processing Error] (B [Math Processing Error] F)

2. [Math Processing Error]F /[Math Processing Error]

3. (B [Math Processing Error] F) 1, simp

4. [Math Processing Error] 2, 3, MT

Let's walk through a few more together, eh?

1. P [Math Processing Error] (R [Math Processing Error] Q)

2. [Math Processing Error] P

3. (Q [Math Processing Error] G) / (R [Math Processing Error] G)

What's a person to do? Remember, **it's a game**: just figure out what you need, look in your toolkit, find the right tool, then use it!

Here, we need to combine 3 with the conditional in 1 so that we can get our conclusion. We know that our *last step* will be to use Hypothetical Syllogism. I like to work backwards like this in my head and then write down the steps going forwards. So we're now (in our heads) trying to figure out how we might get the

“(R *[Math Processing Error]* Q)” all by itself so that we can do an HS. Remember that HS only works when we have both conditionals all by themselves (when the arrows/horsehoes are the main operators of their respective formulas).

Look at 1 and 2. How do we break apart that disjunction?

Oh yeah! Disjunctive Syllogism!

What else do we need for a disjunctive syllogism?

Oh yeah! The negation of the *other disjunct*.

Where is that “*[Math Processing Error]* P”?

Oh yeah! Line 2!

Now we’re ready to do the whole thing going forwards:

1. P *[Math Processing Error]* (R *[Math Processing Error]* Q)

2. *[Math Processing Error]* P

3. (Q *[Math Processing Error]* G) /(R *[Math Processing Error]* G)

4. (R *[Math Processing Error]* Q) 1, 2, DS

5. (R *[Math Processing Error]* G) 3, 4, HS

We’re done! Check out the last line and the conclusion that comes after the slash: there’re exactly the same. Bravo!

Important Lesson:

It doesn’t matter, for any rule of implication, which order the premises come in.

The lines we’re looking for in our derivation (our natural deduction) don’t have to match the lines in the rule of implication in terms of order. They only have to match in terms of ***structure*** *or* ***form****.*

See how 3 and 4 are the premises for our HS, but they don’t match the order of the premises as they’re written in the rule? Who cares? As long as we have the premises we need. We don’t have to worry about their order. *Modus Ponens* premises can look like this:

1. D

2. D *[Math Processing Error]* H

And Disjunctive Syllogism premises can look like this:

1. *[Math Processing Error]*(Z *[Math Processing Error]* K)

2. (Z *[Math Processing Error]* K) *[Math Processing Error]* Q

The only thing that matters is that we’ve matched the form of the premises exactly.

Another important note: the P’s and Q’s in our rules above can stand for *anything*. So *Modus Tollens* premises can look like this:

1. *[Math Processing Error][Math Processing Error]*(D *[Math Processing Error]* F)

2. W *[Math Processing Error]* *[Math Processing Error]* (D *[Math Processing Error]* F)

And Hypothetical Syllogism premises can look like this:

1. *[Math Processing Error][Math Processing Error]*(D *[Math Processing Error]* F) *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)

2. *[Math Processing Error]*W *[Math Processing Error]* *[Math Processing Error][Math Processing Error]* (D *[Math Processing Error]* F)

See how the antecedent of one is equivalent to the consequent of the other? That means we can do a Hypothetical Syllogism.

Let’s try another together. This time I’ll turn the difficulty up to 11.

1. [*[Math Processing Error][Math Processing Error]*(D *[Math Processing Error]* F) *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)]

2. [*[Math Processing Error]*W *[Math Processing Error] [Math Processing Error][Math Processing Error]* (D *[Math Processing Error]* F)]

3. [*[Math Processing Error]*W *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)] *[Math Processing Error]* (*[Math Processing Error][Math Processing Error]*T *[Math Processing Error] [Math Processing Error]*F)

4. *[Math Processing Error][Math Processing Error]*F

5. (G *[Math Processing Error] [Math Processing Error]*T) / *[Math Processing Error]*G

"Woah! That wacky fool thinks we can do THAT!!?!?!??!"

Yes, I do. And I'll thank you not to call me a wacky fool.

"Sorry, I'm just intimidated by the brackets and all of the negations and FIVE premises!?!?! Eeeeek!"

Calm down, dear student, all will be well. We just need to take it one step at a time.

Again, we'll work our way backwards in our head and then work forwards on paper.

Okay, if we're working backwards, we should start with the conclusion, right? We need a "*[Math Processing Error]*G" and looking through the premises, we don't find one. Instead, the only G in the whole problem is in premise 5:

5. *[Math Processing Error]*T) / *[Math Processing Error]*G

Which rule (out of the 4 we've officially learned so far) allows us to grab the antecedent out of a conditional?

You're right! It's *Modus Tollens.* Well done.

Okay, so we'll eventually do an MT to get our conclusion "*[Math Processing Error]*G". What would we need to actually be able to perform the MT?

Ummmmm... I dunno.

Remember, Natural Deduction is one step at a time. Don't get intimidated by trying to think about the whole problem at once. Instead, just focus on the individual step. Here, we're looking at *Modus Tollens* and trying to figure out what would complete the pattern.

[Math Processing Error]

I've replaced the letters with symbols here because sometimes it can be easier to think about replacing a symbol with a complex formula. Here's what we know so far:

The first line is this: (G *[Math Processing Error] [Math Processing Error]*T), which means that the happy face is G and the delta/triangle is...

Uhhhh... T?

Close! But not quite. The first line of MT just deals with ***whatever the antecedent is negations and all*** and ***whatever the consequent is negations and all***. So the delta or triangle symbol is actually *[Math Processing Error]*T. It's whatever comes after the arrow.

So...that would mean *[Math Processing Error]* **is** *[Math Processing Error][Math Processing Error]***T!**

You got it! See? This isn't so scary. Okay, so we need to find *[Math Processing Error][Math Processing Error]*T somewhere in the premises.

Ooooooo! I found it:

3. [*[Math Processing Error]*W *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)] *[Math Processing Error]* (*[Math Processing Error] [Math Processing Error] [Math Processing Error]*F)

Yep! Premise 3 has a *[Math Processing Error][Math Processing Error]*T in its consequent. Let's just focus in on that consequent for a second:

(*[Math Processing Error][Math Processing Error]*T *[Math Processing Error] [Math Processing Error]*F)

How do we break off on disjunct from a disjunction? Disjunctive syllogism! What else do we need to complete a disjunctive syllogism!

Ummm.... I'm stumped again.

A DS has the following form:

[Math Processing Error]

Here, our P is *[Math Processing Error][Math Processing Error]*T and our Q is *[Math Processing Error]*F, which means our *[Math Processing Error]*Q is *[Math Processing Error][Math Processing Error]*F. Following so far?

Yeah, I think so. If we just substitute *[Math Processing Error]***F in for all of the Q's, then we get this, right?**

[Math Processing Error]

Right-o! Well done. And then if we put in *[Math Processing Error][Math Processing Error]*T for the P's, then we get this:

[Math Processing Error]

So, now we're looking for a *[Math Processing Error][Math Processing Error]*F to complete our DS. And it turns out that premise 4 *just is* *[Math Processing Error][Math Processing Error]*F. Bravissimo! So we know we can do our DS when the time comes. Remember, though, that we were looking at the *consequent* of a conditional premise. How do we break the consequent out of a conditional?

Modus Tollens!

Yes! You got it. What do we need to find in order to complete Modus Tollens?

The...Er...Um...Antecedent?

Say it with confidence!

The antecedent!

Now we're getting somewhere. What's the antecedent to premise 3?

This mess?

3. *[Math Processing Error] [Math Processing Error]* (*[Math Processing Error][Math Processing Error]*T *[Math Processing Error] [Math Processing Error]*F)

Yes, the main operator in premise 3 is the conditional that comes after the section you've circled. And whatever comes before the main conditional operator is the antecedent. Looking at that antecedent all by itself:

[*[Math Processing Error]*W *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)]

What sort of formula is this? In other words, what's the main operator?

It's another conditional, right?

Yep! Let's look at premises 1 and 2:

1. [*[Math Processing Error][Math Processing Error]*(D *[Math Processing Error]* F) *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)]

2. [*[Math Processing Error]*W *[Math Processing Error] [Math Processing Error][Math Processing Error]* (D *[Math Processing Error]* F)]

Do you see how that antecedent above is related to premises 1 and 2? It's the antecedent from one premise combined with the consequent from another premise.

1. [*[Math Processing Error]*

2. [*[Math Processing Error]*]

That means we're going to use which of our 4 rules?

Hypothetical syllogism!

Yes! Well done. Now we've gone through the whole problem backwards and we're ready to go through the derivation forwards on paper.

Okay, so I first copy the problem down:

1. [*[Math Processing Error][Math Processing Error]*(D *[Math Processing Error]* F) *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)]

2. [*[Math Processing Error]*W *[Math Processing Error] [Math Processing Error][Math Processing Error]* (D *[Math Processing Error]* F)]

3. [*[Math Processing Error]*W *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)] *[Math Processing Error]* (*[Math Processing Error][Math Processing Error]*T *[Math Processing Error] [Math Processing Error]*F)

4. *[Math Processing Error][Math Processing Error]*F

5. (G *[Math Processing Error] [Math Processing Error]*T) / *[Math Processing Error]*G

And then I work backwards from the order we just discussed. So, the last thing we talked about was doing a HS on 1 and 2:

6. [*[Math Processing Error]*W *[Math Processing Error]* ((Z *[Math Processing Error]* F) *[Math Processing Error]* Q)] 1, 2, HS

Excellent!

Then I do a *Modus Ponens* with 3:

7. (*[Math Processing Error][Math Processing Error]*T *[Math Processing Error] [Math Processing Error]*F) 3, 6, MP

Then I'm ready to do a Disjunctive Syllogism:

8. *[Math Processing Error][Math Processing Error]*T 4, 7, DS

And finally, a *Modus Tollens*! We're so close I can taste it!

9. *[Math Processing Error]*G 5, 8, MT

And we're done!

Very nice! Don't forget to make sure that the last line you've derived is *identical* to the conclusion we were meant to derive.

It is! *[Math Processing Error]***G is the same as** *[Math Processing Error]***G!**

See? Even on this behemoth of a problem with 9 total lines (!!!) you were able to think through it step-by-step and solve the derivation. **Don't be intimidated** by long or complex problems. This is all pretty mechanical and rote, even when it gets complex. *There's no magic special ability called "logical reasoning" that some of us have and some of us don't. YOU CAN DO THIS!*

---

[1] But seriously, some people call this "*modus tollendo ponens*" or "mode of positing by denying".

8.3: More Rules of Implication

Let's introduce the rest of the rules of implication:

> **Constructive Dilemma**
>
> *(a 'Dilemma' is a situation where one must choose between two* ("di") *options ("lemmae"))*
>
> *If I find a* **conjunctive premise** that is a conjunction between two conditionals and a **disjunctive premise** that is a disjunction between both antecedents of those conditionals, then I can write a **disjunctive conclusion** that is a disjunction between both consequents. I justify this by writing "CD."
>
> *[Math Processing Error]*

This is the most complex rule we'll learn, but conceptually speaking it's not all that complicated. It's just two *Modus Ponens*es right next to one another. See how the "F*[Math Processing Error]*G" from the first line and the "F" from the second line make a modus ponens with the "G" in the conclusion? Same for the letters on the right half of each line. The trick in recognizing this one, though, is getting the symbols straight. It's always:

- a conjunction between two conditionals
- a disjunction between the antecedents
- a disjunction between the consequents

There's a lot to say about this, but for now it suffices to say that it's called dilemma for a reason: it's the form of a dilemma that you might find in real life: If we go this way, this consequence will follow; but if we go this other way, this other consequence will follow. These are our only two ways of going, so we're stuck with either of the two consequents.

> **Simplification**
>
> *(It's the* simplest *rule there is!)*
>
> *If I find a conjunction, I can put either conjunct on a line by itself. I justify this move by writing "simp".*
>
> | | |
> |---|---|
> | *[Math Processing Error]* | *[Math Processing Error]* |

Pretty simple right? Need I say more? Just remember that you *actually have to use this rule*. Even though intuitively you should be able to use anything in a conjunction, you must bring it down to its own line using simp before you use it. For instance, this is not yet *Modus Tollens*:

1. ~D *[Math Processing Error]* G

2. T *[Math Processing Error]* D

Okay, moving on to the next rule:

> **Conjunction**
>
> *('Conjunction' is Latin for "a joining together")*
>
> *I'm allowed to join any two lines by a conjunction on a new line. I justify this move by writing "conj".*
>
> | | |
> |---|---|
> | *[Math Processing Error]* | *[Math Processing Error]* |

Pretty simple, still, right? It's just the opposite of Simplification. Just take whatever two lines you like and stick them together with a conjunction. Think about it intuitively: if I tell you that Obama was president in 2012 and I tell you that he was president in 2009, then I've told you that Obama was president in 2012 **and** in 2009.

Here's the next rule, but real quick: *don't confuse this with conjunction.* One uses a conjunction (can you guess which one?) and the other uses a disjunction.

Addition

I can add **whatever I like** *after a disjunction to any premise I like and write the result on a new line. I justify this move by writing "add".*

[Math Processing Error]

Wait, what?!?!?! I knew I was right to call you a Wacky Fool.

I'm no wacky fool and I'm not lying to you here. I promise.

But you can't just add *whatever* you want wherever you want it. This isn't art class, there are rules! Right?

Yes, of course there are rules, silly. This just happens to be one of them! (And there are rules in art class, too)

It's okay to add whatever you want, because when you add a disjunction, you *weaken* the formula. The worry we always have in logic is getting *stronger* propositions from weaker ones for free. There's no free lunch in logic and there is no magic whereby you can get out of a deductive argument more than you put in. That's what Induction (scientific reasoning, for example) is for.

So, when we *weaken* a proposition when moving from the premise to a new version of that proposition in the conclusion, we're all good by the lights of logic.

For example, if I was to tell you that Tiger Woods is a fantastic golf player then you'd be warranted in inferring that **either** Tiger Woods is a fantastic golf player **or** the moon is made of cheese. If we already *know* that one disjunct is true, then there's no worry about the disjunction as a whole being false (all it takes for a disjunction to be true is for at least *one* disjunct to be true). Similarly, if I tell you I teach logic, then you're allowed to conclude that I either teach logic or music. Yes it's true that I either teach logic or music, I teach logic!

Now we have all 8 rules of Implication on the table. I can't emphasize enough the importance of having all of the rules in front of you while you're trying to figure natural deduction problems out. These are the simplified forms of the rules:

8 Rules of Implication:

Table *[Math Processing Error]*: 8 rules of implication.

Modus Ponens MP	*[Math Processing Error]*	***Constructive Dilemma*** CD	*[Math Processing Error]*
Modus Tollens MT	*[Math Processing Error]*	***Simplification*** simp	*[Math Processing Error]*
Disjunctive Syllogism DS	*[Math Processing Error]*	***Conjunction*** conj	*[Math Processing Error]*
Hypothetical Syllogism HS	*[Math Processing Error]*	***Addition*** add	*[Math Processing Error]*

Let's work through a problem or two that involves all 8 rules of Implication together.

1. (P *[Math Processing Error]* R) *[Math Processing Error]* (M *[Math Processing Error]* P)

2. (P *[Math Processing Error]* M) *[Math Processing Error]* (P *[Math Processing Error]* R)

3. P *[Math Processing Error]* M / R *[Math Processing Error]* P

R and P are on the right sides of two horseshoe formulas. Given that, it feels like we might use Modus Ponens at some point to get one or both of those on their own. We might also use Addition since our conclusion is a disjunction. The problem with both of these strategies is that it doesn't look possible to get a P by itself, so it doesn't seem possible to do that Modus Ponens. Without the Modus Ponens, we also won't be able to get our conclusion via Addition. Consequently, it looks like we have many of the basic ingredients of a Constructive Dilemma, so maybe we should start heading in that direction.

An alternative strategy for getting through this problem is identifying the rules we *can* apply and then applying them as we find we're able. That way we don't need to do a lot of strategizing ahead of time and instead can simply trip our way to the solution—a bit easier than having to work through it all in your head.

Right away we see, looking at lines 1-3, that we could do a Hypothetical Syllogism with 1 and 2. We at least have an inkling, though, that we're headed towards a Constructive Dilemma, and that HS doesn't seem like it will get us closer to our goal. So let's skip it.

The other move that is open to us right away is to do an MP with 2 and 3. Let's go ahead an do that, throwing caution to the wind!

4. P *[Math Processing Error]* R 2, 3, MP

What now? It turns out we've opened up another MP by deriving the antecedent of 1:

5. M *[Math Processing Error]* P 1, 4, MP

Okay now we have P implies R and M implies P. Our conclusion is R or P. This is a perfect set up for Constructive Dilemma. We just need to look at the rule for constructive dilemma to help us determine how to construct the premises of the rule.

[Math Processing Error]

It looks like we need the implication (arrow/horseshow) formulas to be joined via conjunction (it's an "*[Math Processing Error]*" above). How can we do that? Well, we just learned a simple rules for making a conjunction: find both conjuncts on their own separate lines and then use the rule "Conjunction" to stick 'em together with a *[Math Processing Error]* or a *[Math Processing Error]* (they both mean the same thing: conjuinction).

6. (P *[Math Processing Error]* R) *[Math Processing Error]* (M *[Math Processing Error]* P) 4, 5, Conj

Great, now we have the first line of Constructive Dilemma above. Where's the second line (F *[Math Processing Error]* P)? Well, in this case, it shoulw be (P *[Math Processing Error]* M). Do we have (P *[Math Processing Error]* M)?

Yeppers, Mr. Teacher Man. That's just line 3 as it is! So we can do CD with 3 and 6!

7. R *[Math Processing Error]* P 3, 6, CD

Double check that your conclusion and your last derived line are identical and if they are, call it a day!

How about we go again?

1. (L *[Math Processing Error]* T) *[Math Processing Error]* (B *[Math Processing Error]* G)

2. L *[Math Processing Error]* (K *[Math Processing Error]* R) / L *[Math Processing Error]* B

It's easy enough to get the L by itself, right? Do you know which rule we'll use?

Yeah, that's pretty clear. Use Simplification to break the L off of that conjunction in 2.

You got it. So smart.

3. L 2, Simp

So the last step will be to use which rule? Conjunction? Yep! We'll take that L in 3 right above and conjoin it together with a B. Where's B?

Up there in Line 1 on the consequent side.

Cool, and it's in a conjunction, so we know we can simplify again once we break that consequent out of the horseshoe in 1. How are we going to get that consequent by itself?

To take the right side out of an arrow/horseshoe, you use....... Modus Ponens!

Perfecto. So we have to build (L *[Math Processing Error]* T). How do we get that?

Dunno dude, we only have an L.

Maybe if we could...Add to that L?

Oh yeah! Duh. Addition.

4. L *[Math Processing Error]* T 3, Add

Yep. It's a bit counterintuitive, but since (L *[Math Processing Error]* T) is weaker than L—claiming that *either I'm going to be president or I'm going to order a salad for lunch* is obviously not as strong a claim as to claim that *I'm going to be president!*—this is logically valid.

Okay, now that I've got the left side of 1, I can do...Modus Ponens!

5. B *[Math Processing Error]* G 1, 4, MP

And I wanted that B. That's the reason we've been doing all of this Addition and Modus Ponens nonsense. Just simplify.

6. B 5, Simp

And then, as we decided earlier, I can use Conjunction to stick 'em together and get my conclusion. Double check: is it the same as the conclusion after the slash above?

7. L *[Math Processing Error]* B 3, 6, Conj

8.4: Rules of Replacement

The final rules are fairly simple (in fact, they're *simpler* than the rules of implication), but work in a very different way than the rules of implication. They're called the rules of *replacement* because they allow you to simply transform or replace a formula (or subformula) with a logically equivalent formula.

If you look above at the rules of implication, you'll notice that they have a sort of "premise-conclusion" format. If you find the premises, then you can derive the conclusion.

Rules of replacement have a different structure because each is simply a statement that two formulas are logically equivalent. So instead of going in one direction, we can move in **either direction** in the rule. Here's an almost ridiculously simple rule of replacement:

$$\sim\sim P :: P$$

The first thing to note is that little square of dots. That's what's called a "metalogical" symbol, but you don't have to worry too much about that. What it means for our purposes is simple: you can replace these with each other (i.e.[1] they are logically equivalent). A :: B means "you can replace A with B and B with A. The first difference between rules of replacement and rules of implication, therefore, is that rules of replacement are **bidirectional**—one can move from the left to the right or from the right to left. With the rules of implication, we were only allowed to move *down* from the premises to the conclusion—we can't move *up* from the conclusion back to the premises. That wouldn't be a valid move. The rules of replacement work in both ways.

This means we can derive from *any formula* the double-negated form of that formula and vice versa. Put another way, we can *remove two negations* or add two negations to any formula we want. See how simple some of these rules can be?

Let's introduce the rest of the rules, and then talk a bit more about differences between rules of replacement and rules of implication. I use Greek Letters *Delta*, *Ohm*, and *Phi* to help get us into the habit of thinking of these as *general patterns* rather than as specific formulas with just single letters in place of the disjuncts, conjuncts, antecedents, consequents, etc.

9. DeMorgen's Rules (DM):

[Math Processing Error]

[Math Processing Error]

In short: "distributing" a negation the way one might distribute an exponent in math (or, for that matter "undsitributing" as one does moving from right to left in these rules) results in flipping an "and" to an "or" and vice versa.

10. Commutativity (Com):

[Math Processing Error]

[Math Processing Error]

In short: you can switch conjuncts with one another and disjuncts with one another. Note: only works with conjunction (*[Math Processing Error]* or *[Math Processing Error]*) and disjunction (*[Math Processing Error]*). It doesn't work with any other connectives.

Associativity (Assoc)

[Math Processing Error]

[Math Processing Error]

In short: if you have two adjacent disjunctions or conjunctions, you may move the parentheses to *associate* the two letters that are not being associated in the base formula. Note: only works with conjunction (*[Math Processing Error]* or *[Math Processing Error]*) and disjunction (*[Math Processing Error]*). It doesn't work with any other connectives.

12. Double Negation (DN)

[Math Processing Error]

In short: you can add or remove two **immediately adjacent** negations to a formula or subformula. Note: negations cannot be separated, so this doesn't work: ~(~Z*[Math Processing Error]*X) // (Z*[Math Processing Error]*X) [invalid!]

13. Transposition (Trans)

[Math Processing Error]

In short: you can switch antecedent and consequent as long as you add one negation to each or take away one negation from each. Note: must be combined with DN if you want to add one negation and take one negation away using Trans.

14. Material Implication (Impl)

[Math Processing Error]

In short: you can change an implication into a disjunction as long as you negate the antecedent. You can also replace a disjunction with at least one negated formula with an implication, so long as the left disjunct is negated. If you do, removed *one* negation from the left disjunct.

Another important difference between rules of replacement and rules of implication is that rules of replacement can be used on *subformulas*. A subformula is any formula that is part of another formula. "(A*[Math Processing Error]*B)" is a formula, so it's a subformula of #1 below. "P *[Math Processing Error]* *[Math Processing Error]*" is not a well-formed formula, so it's not a subformula of #3 below. I can apply a rule of replacement to only *part* of a formula. So the following derivation is deductively valid even though all of the changes happen to only part of the premise and subsequent formulas:

1. P *[Math Processing Error]* (A*[Math Processing Error]*B) / P *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* *[Math Processing Error]*B)

2. P *[Math Processing Error]* (*[Math Processing Error]*A*[Math Processing Error]*B) 1, Impl

3. P *[Math Processing Error]* *[Math Processing Error]**[Math Processing Error]*(*[Math Processing Error]*A *[Math Processing Error]* B) 2, DN

4. P *[Math Processing Error]* *[Math Processing Error]*(*[Math Processing Error]**[Math Processing Error]*A *[Math Processing Error]**[Math Processing Error]*B) 3, DM

5. P *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* *[Math Processing Error]*B) 4, DN

If we know that each of the replaced subformulas are logically equivalent, then (since equivalence is transitive), we also know that all of them will be materially equivalent (biconditional) to P (if, that is, proposition 1 is true).

To summarize: rules of replacement are **bidirectional** and can be applied to **subformulas** (parts of formulas); whereas rules of implication are **unidirectional** and apply only to **whole Well-Formed Formulas/WFFs.**

Stacking rules

Let's discuss another revision: your instructor may be okay with you stacking a couple of rules on one line. For instance, you may be able to throw a Trans and DN together to add a negation to one part and remove one from another. The important caveat is that you must cite the rules in order to minimize confusion. The above proof might, for instance, look like this:

1. P *[Math Processing Error]* (A*[Math Processing Error]*B) / P *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* *[Math Processing Error]*B)

2. P *[Math Processing Error]* *[Math Processing Error]**[Math Processing Error]*(*[Math Processing Error]*A *[Math Processing Error]* B) 1, Impl, DN

3. P *[Math Processing Error]* *[Math Processing Error]*(A *[Math Processing Error]* *[Math Processing Error]*B) 2, DM, DN

Generally, this is restricted to Com and DN since these are quick and simple changes rather than dramatic transformations of the formulas involved.

Com and the Strictness of Rules

Once we have these rules on our tool belts, we can actually go back and be a bit stricter with some of the previous rules. Sometimes, your instructor may restrict the rules to apply only to a very specific pattern. Whereas other instructors will be more liberal with the rules. For instance, DS looks like this:

[Math Processing Error]

Notice how it's the *left* disjunct being negated here. If we left the rule at that, we'd have to use Commutativity, if they were reversed, to switch around the P and Q before we do DS. So it'd have to look like this:

1. Q *[Math Processing Error]* P

2. *[Math Processing Error]*P /Q

3. P *[Math Processing Error]* Q 1, Com

4. Q 2,3 DS

In this case if you stacked rules, you'd need to cite "Com" as the first rule because DS requires Com to happen *first*.

Similarly with simplification, which sometimes is written to only allow you to simplify the *left* conjunct. Here's an example of a proof that follows this rule strictly:

1. P *[Math Processing Error]* ~Q

2. R *[Math Processing Error]* Q / ~R *[Math Processing Error]* P

3. P 1, Simp

4. ~Q 1, Com, Simp

5. ~R *[Math Processing Error]* Q 2, Impl

6. Q *[Math Processing Error]* ~R 5, Com

7. ~R 4, 6, DS

8. ~R *[Math Processing Error]* P 3, 7, Conj

All 14 Rules (simplified form)

Table *[Math Processing Error]*: All 14 rules (simplified form).

Modus Ponens MP	*[Math Processing Error]*	***Constructive Dilemma*** CD	*[Math Processing Error]*
Modus Tollens MT	*[Math Processing Error]*	***Simplification*** simp	*[Math Processing Error]*
Disjunctive Syllogism DS	*[Math Processing Error]*	***Conjunction*** conj	*[Math Processing Error]*
Hypothetical Syllogism HS	*[Math Processing Error]*	***Addition*** add	*[Math Processing Error]*
DeMorgan's Rules DM	*[Math Processing Error]* *[Math Processing Error]*	**Double Negation** DN	*[Math Processing Error]*
Commutativity Com	*[Math Processing Error]* *[Math Processing Error]*	**Transposition** Trans	*[Math Processing Error]*
Associativity (Assoc)	*[Math Processing Error]* *[Math Processing Error]*	**Material Implication** Impl	*[Math Processing Error]*

All 14 Rules (in simplified phrases, not strictly worded)

Table *[Math Processing Error]*: All 14 rules (in simplified phrases, not strictly worded).

Modus Ponens MP	Find the antecedent, write the consequent	***Constructive Dilemma*** CD	Like two MPs side-by-side
Modus Tollens MT	Negate the consequent, negate the antecedent	***Simplification*** simp	Write a conjunct by itself
Disjunctive Syllogism DS	Negate one disjunct, write the other	***Conjunction*** conj	Stick two formulas together with a '*[Math Processing Error]*' (or '*[Math Processing Error]*') between
Hypothetical Syllogism HS	Take out the "middle man"	***Addition*** add	Add anything to the right of a '*[Math Processing Error]*'
DeMorgan's Rules DM	Distribute or un-distribute a negation, change *[Math Processing Error]* to *[Math Processing Error]* and vice-versa	**Double Negation** DN	Add or remove two negations
Commutativity Com	Switch conjuncts or disjuncts with their partners	**Transposition** Trans	Like MT but in replacement form
Associativity (Assoc)	Move parentheses around two *[Math Processing Error]*'s or two *[Math Processing Error]*'s (or *[Math Processing Error]*'s).	**Material Implication** Impl	Switch between *[Math Processing Error]* and *[Math Processing Error]*, negate or un-negate the antecedent

Now let's do a couple of proofs involving the rules of replacement.

1. ~(A*[Math Processing Error]*~B)

2. ~(B*[Math Processing Error]*~C) / A *[Math Processing Error]* C

If we pay attention to the placement of the A, the B's, and the C in 1 and 2, we might notice that they're pretty close to a Hypothetical Syllogism. There's a B on the right of one and on the left of the other, so there's a "middle man" that we can eliminate using HS. The result would be the conclusion. So our goal is to get 1 and 2 into the proper format to be able to apply HS.

How do we do that? Well, looking at 1 and 2, we might notice that there's really only one rule which allows us to transform a negated conjunction. What rule is that??? Look at the table of rules above. I'll wait. Right-O! DeMorgen's rules are the only way we know of (in this class) to transform a negated conjunction or disjunction. So let's just go ahead and apply DM and see what happens. We distribute the outside negation to both conjuncts and then we change the dot to a vee like so:

3. ~A *[Math Processing Error]* ~~B 1,DM

4. ~B *[Math Processing Error]* ~~C 2,DM

Remember we want to make conditionals. What rule allows us to change a disjunction into a conditional? Yep! Material Implication will do just that. We also don't want all of those ugly negations, so we just apply double negation to remove them:

5. A *[Math Processing Error]* B 3, DN, Impl

6. B *[Math Processing Error]* C 4, DN, Impl

What was that last step we decided on waaaaay at the beginning of the proof? Right on. Hypothetical Syllogism to take of the middle man "B" and leave us with our conclusion:

7. A *[Math Processing Error]* C 5, 6, HS

QED! (QED stands for the Latin "Quod Erat Demostrandum" or "which is what was to be demonstrated". So we can say "QED" when we're done with a proof or demonstration—that is, when we've reached the conclusion—as a way of saying "I proved what I was supposed to prove!")

How about we try another?

Thank you, Sir! May I please have another!

Uh oh, it's bad news when your students start talking to you like a Private in Boot Camp! Anyways, here's the problem:

1. S *[Math Processing Error]* ~M

2. (P *[Math Processing Error]* R) *[Math Processing Error]* M / ~P *[Math Processing Error]* ~R

One way to think about a problem that looks like this is that we have to find ~P and then find ~R and then we can use Conjunction to stick them together to make the conclusion. This won't really work in this case, though, since the P and R are right next to one another in the same subformula, so we won't really be able to get them alone—at least not without some heartache. With that in mind, let's try something different. What if we tried to negate (P *[Math Processing Error]* R)? If we did that, we could use a certain rule to turn ~(P *[Math Processing Error]* R) into (~P *[Math Processing Error]* ~R). Which rule?

DeMorgen's Rule!

Yepperino. So now all we have to do is negate the left side of the condition #2. How do we negate the left side of a conditional?

Modus Ponens!

Close, but not quite.

Modus Tollens!

There you go. Remember "Ponens" means affirming and "Tollens" means negating. In this case we want to negate. What else do we need in order to do an MT?

~M

Yes, dear student. You're being so smart. Now how are we going to get that ~M? To finish the proof out, I'm going to use a stricter version of Simplification according to which one can *only simplify the left side* of a conjunction. If that's the case, we'll need to switch #1 around left to right before simplifying. Which rule allows us to do this? Com:

3. ~M *[Math Processing Error]* S 1, Com

Excellent, now simplify and then Modus Tollens:

4. ~M 3, Simp

5. ~(P *[Math Processing Error]* R) 2,4,MT

What happens next? Remember way back to the beginning when we were talking about turning 5 into the conclusion. We'll use DeMorgen's Rule:

6. ~P *[Math Processing Error]* ~R 6, DM

---

[1] Fun Fact: i.e. is an abbreviation of "id est", which is Latin for "that is". 'e.g.' is an abbreviation of 'exempli gratia', which is Latin for "for example".

8.5: More Rules

Here are a few more rules to flesh out our natural deduction system. Your instructor may or may not want you to devote the time and energy to learning them.

The first rule fills a gap in our current system: we have no way of dealing with biconditionals! That's a pretty glaring omission. Without rules like this one, we're stuck treating biconditionals as unbreakable units or atoms that are more or less useless aside from what they can do as a whole. We've used biconditionals before, but we've never looked *inside* of a biconditional while doing natural deduction. This rule fixes that problem. Treat it as a *rule of replacement*:

Biconditional (Bc):

[Math Processing Error]

You can replace any biconditional statement with a conjunction of both corresponding conditional statements. After using Biconditional (a sort of combination of "biconditional introduction" and "biconditional elimination"), you can simplify the conjunction to get a single direction of the conditional. You might be able to see intuitively that we could do biconditional modus ponens and biconditional modus tollens. Some more powerful systems have these as rules. We won't make a separate rule for them here, but they are perfectly valid rules that simply bypass the steps of using Biconditional and then Simplification.

Here's a second biconditional rule you may want to make use of. It's a bit more counterintuitive in that it rests on the fact that it's a *material* biconditional and material biconditionals are sort of weird:

Bc2:

[Math Processing Error]

Also, Associativity works for biconditionals too:

Associativity:

[Math Processing Error]

There are a few more sort of standard rules that your instructor may want to include:

Exportation:

[Math Processing Error]

Distribution:

[Math Processing Error]

[Math Processing Error]

Negation of Conditional:

[Math Processing Error]

Negation of Biconditional:

[Math Processing Error]

You might also find useful certain sorts of axiomatic or trivial rules. Here are a few that your instructor might want to make use of:

Trivial Implication:

[Math Processing Error]

Excluded Middle MP:

[Math Processing Error]

Tautology 1 (T1):

Notice that you don't need any premise to posit a tautology. The justification would be "T1".

[Math Processing Error]

Tautology 2 (T2):

Notice that you don't need any premise to posit a tautology. The justification would be "T2".

[Math Processing Error]

Tautology 3 (T3):

[Math Processing Error]

Or

[Math Processing Error]

8.6: Why Learn Natural Deduction?

As per usual, the reason for learning these various procedures and concepts in logic can be a bit mysterious. Why is it important that we know this? It was already unlikely that we'd build truth tables to figure out whether or not an argument we find "in the wild" is valid or invalid, or to test the formal consistency of our beliefs. Much less will we bust out our inference rules and *prove* that an argument is valid, given that we often must already *know* it's valid if we're going to commit to spending the time to do a natural deduction proof!

This is, in fact, true for most things you learn in college. You won't regularly calculate the standard deviation of a distribution by hand like you do in Statistics. You won't regularly perform integral calculus, recite the stages of the hero's journey, identify auxiliary verbs, or conjugate the past imperfect passive in a different language. The point of learning all of these things isn't that they will be immediately useful for your everyday life. You'll pick up all of those useful skills on your own by simply living your life.

The point is not to give you a tool that you'll regularly use in your everyday life. Much of what you learn in school is about giving you *understanding* rather than immediately useful knowledge. This is one more item in this general theme: the point is to understand deductive inferences at an intuitive level rather than to build out a toolkit that one can put to use on a Reddit discussion board or on an assembly line.

If you want to understand how a car works beyond "if I push that pedal down it goes forward," you have to look under the hood and get your hands dirty in order to see how things fit together. The same goes for Natural Deduction: if you want to know how logical reasoning works, it is best to "look under the hood." The inference rules are these simple components of reasoning that we understand at an intuitive level, so if we find out that a complex inference is in fact a combination of six of these simple rules, then we've found out how that complex inference works and we now understand something quite complex at an intuitive level.

A few reasons for learning Natural Deduction are that we can use it to understand more clearly:

1. *How* premises are related to their conclusions
2. What a complex inference is built out of and what it means for complex inferences to be built out of simpler inferences.
3. What it means for a premise to be irrelevant to an inference
4. How deductive inference works at a more atomic level

In the process of learning Natural Deduction, moreover, one cannot help but internalize some of the inference rules we've learned. The more you internalize these rules, the better sense you have of *what implies what*. When you've been doing logic as long as I have, you can pretty quickly see that "if you don't go out with me, then I'll be distraught" deductively implies "either you go out with me or I'll be distraught". With a little more thought, I can figure that it also implies "it won't happen that both you won't go out with me and I won't be distraught" or "if I'm not distraught, then it means that you agreed to go out with me.' The more logic we do, the easier these things are to recognize.

8.7: Conversational Implicature

Now that we've talked a lot about logical implication, let's take a second to talk about a perhaps more important phenomenon: conversational implicature.

If I go into the kitchen, find the main course missing from the counter, and ask "what happened to the roast?" You might, noticing that the dog looks satisfied and full, say "well, the dog certainly looks happy." Logically, you haven't answered the question. We can say a lot about the *logical* implications of what you've said, but that won't tell us what you've accomplished by saying what you said. Instead, we need the concept of a conversational implicature: a way we can communicate more than is logically implied by what we've said by relying on the rules of conversation.

In this example, there's a rule in conversations that says "don't say things that aren't relevant to the topic at hand." I'll therefore assume that you mean "the dog looks happy" to be a relevant piece of information to my question. So it seems like what you're *implicating* (even though you're not logically implying it) is that the dog ate the roast.

To highlight some of the awkwardness that can happen when we accidentally implicate things, here is a list of funny implicatures and the rules they violate which give them the unintended meaning:

- I have a brother, I also have another brother
  - *This is called "scalar implicature": it's where you say something and everyone assumes that you've said the most informative quantitative statement you could have. Here, it's assumed that you would have made the strongest claim available: I have two brothers*
- Mitch Hedberg has a great one liner that's also a sort of scalar implicature: "I used to do drugs. I still do, but I used to, too."
- Another Mitch Hedberg gem (he's a master of implicatures): "I want to hang a map of the world in my house. Then I'm gonna put pins into all the locations I've traveled to, but first I'm gonna have to travel to the top two corners of the map, so it won't fall down."
  - *The rule is something like "don't say you're going to do something like hang something on your wall if you don't mean to say that you're going to do everything required to do it in a normal way." By ignoring this rule, Mitch makes us assume that he is going to hang the map normally and then plays with that assumption.*
- More Mitch Hedberg: "I wrote a script and I gave it to a guy who reads scripts. And he read it and he said he really likes it, but he thinks I need to re-write it. I said "Screw that, I'll just make a copy.""
  - *Of course, we're assuming that the editor in question meant that the script needed serious revision because of a rule that says something like "don't say something needs to be rewritten when you just mean it needs to be copied." or more generally "make appropriate requests given the technology of the time." Mitch is playing on that assumption by acting like he interpreted the editor the opposite way from the way he intended.*
- Mitch Hedberg again: This one commercial said "Forget everything you know about slip covers," so I did. And it was a load off my mind. Then the commercial tried to sell slip covers, but I didn't know what they were.
  - *We all know that when someone says "forget everything you know about x" they aren't intending the phrase literally and instead, by convention, mean something like "here's a paradigm-shattering new product," or "I'm about to show you something which defies your previous assumptions about slip covers." Mitch is instead interpreting it literally, which isn't what convention dictates, even if Logic might dictate that we do so (in this case, it's unclear what logic would tell us to do)*

To take a more relevant and controversial example: If my friend says, "Black people matter," and I respond with, "All people matter", then I'm saying something which *logically implies* that "black people matter." But even though my statement has this logical implication, there's a rule of conversation that says something like "don't make a broader statement when you mean to communicate something more specific." In breaking this rule, the *conversational implicature* is that people shouldn't be claiming that "black people matter" and instead should be claiming that "all people matter." The implicature is that there's something somehow wrong with saying "Black people matter" and that they should be saying something more general—that the person affirming the worth of black people has made some sort of mistake. In saying "all people matter" in response to someone saying, "black people matter", therefore, I'm doing something equivalent to the following:

"I love Oak trees"

"Well, I love all trees"

The implicature is that there's something wrong with saying one loves oak trees. That there's something mistaken in making the more specific claim. Again, it's like the following:

"We need to cure cancer as soon as possible."

"We need to cure all diseases as soon as possible."

The second speaker, because they are "zooming out" to a less specific claim, is implicating that there's a mistake in the first claim—that it's *too* specific. But presumably there's no problem with saying we need to cure cancer as soon as possible. This is partially because cancer is a particularly widespread and deadly disease.

Conversely, some people read the saying "black lives matter" as having the *implicature* that "black lives matter and other lives do not," or "black lives matter more than other lives." The idea here, I take it, is that in making this claim we are singling out a specific group when what we mean to be doing is affirming the importance of all lives and so we are making a mistake: we're implicating that black lives exclusively or especially matter. I don't think there is a mistake in saying "black lives matter," but the goal here is to understand how confusions begin (apart, perhaps, from problematic racial attitudes).

Both of these examples show how communication—particularly about sensitive and important issues—breaks down easily. When we're not all in agreement on the rules of the conversation and the common ground stock of knowledge we share, we can easily make claims that get misinterpreted. We all need to be charitable, take care and take our time, and be humble in our verbal interactions with one another. As always, try to focus more on your own mistakes than those of others.

8.E: Chapter Seven (Exercises)

Exercise $8.E.1$: Identifying Instances of the Rules

Identify which (single) rule is being exemplified in each inference. Remember that the underline separates premises from their conclusion.

A. $\dfrac{[(A \wedge \neg B) \rightarrow (P \vee Q)]}{[\neg(P \vee Q) \rightarrow \neg(A \wedge \neg B)]}$

B. $\dfrac{\begin{array}{l}[(P \vee Q) \vee \neg(A \wedge \neg B)] \\ \neg\neg(A \wedge \neg B)\end{array}}{(P \vee Q)}$

C. $\dfrac{[((Z \vee X) \rightarrow Y) \wedge (Y \leftrightarrow W)]}{((Z \vee X) \rightarrow Y)}$

D. $\dfrac{\begin{array}{l}[((Z \vee X) \rightarrow Y) \rightarrow (Y \leftrightarrow W)] \\ ((Z \vee X) \rightarrow Y)\end{array}}{(Y \leftrightarrow W)}$

E. $\dfrac{\begin{array}{l}[\neg B \rightarrow (P \vee Q)] \\ \neg(P \vee Q)\end{array}}{\neg\neg B}$

F. $\dfrac{\begin{array}{l}[((Z \vee X) \rightarrow Y) \vee (Y \leftrightarrow W)] \\ [(((Z \vee X) \rightarrow Y) \rightarrow T) \wedge ((Y \leftrightarrow W) \rightarrow \neg(W \leftrightarrow Z))]\end{array}}{[T \vee \neg(W \leftrightarrow Z)]}$

G. $\dfrac{(((Z \vee X) \vee Y) \vee T)}{(Z \vee (X \vee Y)) \vee T)}$

H. $\dfrac{[\neg(P \vee Q) \vee (A \leftrightarrow \neg B)]}{[(P \vee Q) \rightarrow (A \leftrightarrow \neg B)]}$

I. $\dfrac{[(P \vee Q) \rightarrow (A \leftrightarrow \neg B)]}{[\neg(P \vee Q) \vee (A \leftrightarrow \neg B)]}$

J. $\dfrac{(((Z \vee X) \rightarrow Y) \rightarrow T)}{(((Z \vee X) \rightarrow Y) \rightarrow T) \vee [((Y \leftrightarrow W) \rightarrow \neg(W \leftrightarrow Z))]}$

K. $\dfrac{\begin{array}{l}[((Y \leftrightarrow W) \rightarrow \neg(W \leftrightarrow Z)) \rightarrow (T \vee \neg(W \leftrightarrow Z))] \\ ((Y \leftrightarrow W) \rightarrow \neg(W \leftrightarrow Z))\end{array}}{(T \vee \neg(W \leftrightarrow Z))}$

L. $\dfrac{\begin{array}{l}(((Z \vee X) \rightarrow Y) \rightarrow T) \\ (T \rightarrow \neg(W \leftrightarrow Z))\end{array}}{[((Z \vee X) \rightarrow Y) \rightarrow \neg(W \leftrightarrow Z)]}$

Exercise $8.E.2$: Using the First 4 Rules

Derive the conclusion (after the slash) from the numbered premises using only the first 4 rules we learned in section 7.2.

A.

- 1. ~C⊃ (A⊃C)
- 2. ~C /~A

B.

- 1. ~W⊃ [~W⊃ (X⊃W)]
- 2. ~W / ~X

C.

- 1. (P → Q)
- 2. (Q → R)
- 3. (R → S)
- 4. (S → T) / (P → T)

D.

- 1. (((Z ∨ X) → Y) → T)
- 2. (T → ¬ (W ↔ Z))
- 3. ((Z ∨ X) → W)
- 4. (W → Y) / ¬ (W ↔ Z

E.

- 1. ~M ∨ (B ∨ ~T)
- 2. (B ⊃ W)
- 3. ~~M
- 4. ~W / ~T

F.

- 1. (~S ⊃ D)
- 2. [~S ∨ (~D ⊃ K)]
- 3. ~D / K

G.

- 1. [A ⊃ (E ⊃ ~F)]
- 2. [H ∨ (~F ⊃ M)]
- 3. A
- 4. ~H / (E ⊃ M)

H.

- 1. [G ⊃ [~O ⊃ (G ⊃ D)]]
- 2. (O ∨ G)
- 3. ~O / D

I.

- 1. P⊃ (G⊃T)
- 2. Q⊃ (T⊃E)
- 3. P
- 4. Q / (G⊃E)

J.

- 1. (~S ⊃ D)
- 2. ~S ∨ (~D⊃ K)
- 3. ~D /K

K.

- 1. X → (Y → Z)
- 2. X → (Z → W)
- 3. (T ∨ X)
- 4. ~T / (Y → W)

L.

- 1. X ∨ (Y ∧ Z)
- 2. (Y ∧ Z) → W
- 3. (X → T)
- 4. ¬ T /W

? Exercise $8.E.3$: Using the First 8 Rules

Derive the conclusion (after the slash) from the numbered premises using only the first 8 rules (the rules of implication) we learned in sections 7.2 and 7.3.

A.

- 1. (~A → H)
- 2. (R → ~B)
- 3. (~A ∨ R) / (H ∨ ~B)

B.

- 1. E ⊃ (A • C)
- 2. A ⊃ (F • E)
- 3. E / F

C.

- 1. (~F ∨ M) ⊃ (P ∨ B)
- 2. (F ⊃ P)
- 3. ~P / B

D.

- 1. M ⊃ (F•G)
- 2. (F ⊃ K)
- 3. W
- 4. (W ⊃ M) / K

E.

- 1. (M⊃F) • (Z⊃W)
- 2. (K • L) • A
- 3. K⊃ (M ∨ Z) / (F ∨ W)

F.

- 1. (M⊃F) • L
- 2. (F⊃G) • A
- 3. (M⊃G) ⊃ [(M⊃F) ⊃W] / W

G.

- 1. (M • F) ∨ (G • W)
- 2. (M • F) ⊃L
- 3. (~L • A)
- 4. G ⊃ (N • O) / N

H.

- 1. (F • A) ⊃ (G • K)
- 2. (M⊃F) • G
- 3. (M • L)
- 4. (M⊃F) ⊃ A / (G ∨ W)

I.

- 1. (M ⊃ P)

- 2. (M • R)
- 3. P ⊃ (Q • S)
- 4. (P•Q) ⊃ (P≡Q) / (P≡Q)

J.

- 1. A ⊃ (B⊃C)
- 2. (A • M)
- 3. B • (F•G) / (C • A)

Exercise $8.E.4$: Using all of the Rules of Implication and Replacement

Derive the conclusion (after the slash) from the numbered premises using any rule we've learned in this text.

A.

- 1. (X ⊃ ~B)
- 2. (D ∨ X)
- 3. B /D

B.

- 1. [Q ∨ (A ∨ C)]
- 2. ¬ C / (A ∨ Q)

C.

- 1. (¬ M⊃B) • (¬ N⊃Q)
- 2. ¬ (M • N) / [(B ∨ Q) ∨ ¬ Z]

D.

- 1. (~X ∨ Y)
- 2. (~Y ∨ Z)
- 3. (X • W) /(X•Y) • Z

E.

- 1. (~X ⊃ T)
- 2. (W • ~T)
- 3. (X ∨ Y) ⊃ Z / (W•Z)

F.

- 1. (X • ~Z)
- 2. (Y ∨ X) ⊃ ~W /~(Z ∨ W)

G.

- 1. ~(X ∨ ~X) / Y

H.

- 1. (G ⊃ E)
- 2. (H ⊃ ~E) / (G ⊃ ~H)

I.

- 1. (~N ∨ B)
- 2. (N ⊃ B) ⊃ T / T

J.

- 1. ~X
- 2. ~(B • Q)
- 3. (~X ⊃ B) / ~Q

K.

- 1. (Y • Z) ⊃ X
- 2. (~X • ~Z)
- 3. Y ⊃ (Y • Z) / (Z ∨ ~Y)

L.

- 1. (F ⊃ ~J)
- 2. H ⊃ (F ∨ G)
- 3. (G ⊃ ~K)
- 4. H / ~(J•K) • (H ∨ F)

M.

- 1. (A ⊃ ~~B)
- 2. A ∨ (~C ∨ ~E)
- 3. ~B
- 4. D ⊃ (C • E) / ~D

N.

- 1. (A ∨ M)
- 2. (A ⊃ M) • (A ⊃ O)
- 3. (M ⊃ O)
- 4. (M ⊃ B) • (O ⊃ A) / (O ∨ B)

CHAPTER OVERVIEW

9: Inductive Reasoning - hypothetical, causal, statistical, and others

Remember way back in Chapter One we introduced a distinction between **deductive** and **inductive** reasoning? Well, we've spent a lot of time on deductive reasoning, so we should spend at least a bit talking about inductive reasoning. Now, informal fallacies often have to do with inductive reasoning, and argument maps are a way of diagramming both inductive and deductive arguments, so we haven't exactly skipped over inductive reasoning, but there are a few things to discuss that deal most specifically with a few different sorts of inductive arguments.

Scientists engage in a particular kind of reasoning to get from the evidence they collect to a theory about how the world works. This type of reasoning is "ampliative," meaning the theory that the sun *always* rises in the morning in the East and sets in the evening in the West is based on seeing a relatively small number of sunrises and sunsets. Obviously, no one has seen *all* of the sunrises that ever happened. So how do we know? We *add* to the evidence by inferring conclusions that go *beyond* the evidence they have.

We everyday folks *also* engage in this kind of reasoning all the time. I get in my car, turn my car, it doesn't start, and I begin to hypothesize what might be the cause. I notice that most of the men I talk to have deeper voices than most of the women I talk to and I (perhaps incorrectly?) infer that men tend to have deeper voices than women. I get cold, put my jacket on, and feel warmer. What do I infer? That the jacket keeps me warm.

What these everyday examples have in common with scientific reasoning is that they, too, are *ampliative*—the conclusions are stronger claims than the premises or evidence. If we couldn't make inductive inferences, we'd be stuck in a strange situation. I'd see the sun rise and then learn…that the sun just rose. I wouldn't learn anything over and above that particular fact. I might put my jacket on when I'm cold, get warmer, and learn…that I put my jacket on this one time and got warmer this one time. Not very exciting conclusions. We need induction to get conclusions that go beyond our evidence.

This chapter will cover a series of types of induction. First, we'll cover **Hypothetical Reasoning**, which is the kind of reasoning that people call the "scientific method." After that, we'll cover the basics of **Causal Reasoning**, **Statistical Generalization**, and **Arguing from Analogy**. We'll also explore the logic of each a bit and look at some pitfalls of each kind of reasoning.

9.1: Hypothetical Reasoning

Suppose I'm going on a picnic and I'm only selecting items that fit a certain rule. You want to find out what rule I'm using, so you offer up some guesses at items I might want to bring:

A Banana

An Egg Salad Sandwich

A grape soda

Suppose now that I tell you that I'm okay with the first two, but I won't bring the third. Your next step is interesting: you look at the first two, figure out what they have in common, and then you *take a guess* at the rule I'm using. In other words, you posit a hypothesis. You say something like

Do you only want to bring things that are yellow or tan?

Notice how at this point your hypothesis goes *way* beyond the evidence. Bananas and egg salad sandwiches have so much more in common than being yellow/tan objects. This is how hypothetical reasoning works: you look at the evidence, add a hypothesis that makes sense of that evidence (one among many hypotheses available), and then check to be sure that your hypothesis continues to make sense of new evidence as it is collected.

Suppose I now tell you that you haven't guessed the right rule. So, you might throw out some more objects:

Grapes

A key lime pie

A jug of orange juice

I then tell you that the first two are okay, but again the last item is not going with me on this picnic.

It's solid items! Solid items are okay, but liquid items are not.

Again, not quite. Try another set of items. You are still convinced that it has to do with the soda and the juice being liquid, so you try out an interesting tactic:

An ice cube

Some liquid water

Some water Vapor

The first and last items are okay, but not the middle one. Now you think you've got me. You guess that the rule is "anything but liquids," but I refuse to tell you whether you got it right. You're pretty confident at this point, but perhaps you're not *certain*. In principle, there could always be more evidence that upsets your hypothesis. I might say that the ocean is okay but a fresh water lake isn't, and that would be very confusing for you. You'll never be quite certain that you've guessed my rule correctly because it's always in principle possible that I have a super complex rule that is more complex than your hypothesis.

So in hypothetical reasoning what we're doing is making a leap from the evidence we have available to the rule or principle or theory which explains that evidence. The hypothesis is the link between the two. We have some finite evidence available to us, and we hypothesize an explanation. The explanation we posit either is or is not the true explanation, and so we're using the hypothesis as a bridge to get onto the true explanation of what is happening in the world.

The hypothetical method has four stages. Let's illustrate each with an example. You are investigating a murder and have collected a lot of evidence but do not yet have a guess as to who the killer might be.

1. The occurrence of a problem

Example *[Math Processing Error]*

Someone has been murdered and we need to find out who the killer is so that we might bring them to justice.

2. Formulating a hypothesis

> ✔ Example [Math Processing Error]
>
> After collecting some evidence, you weigh the reasons in favor of thinking that each suspect is indeed the murderer, and you decide that the spouse is responsible.

3. Drawing implications from the hypothesis

> ✔ Example [Math Processing Error]
>
> If the spouse was the murderer, then a number of things follow. The spouse must have a weak alibi or their alibi must rest on some falsehood. There is likely to be some evidence on their property or among their belongings that links the spouse to the murder. The spouse likely had motive. etc., etc., etc.

We can go on for ages, but the basic point is that once we've got an idea of what the explanation for the murder is (in this case, the hypothesis is that the spouse murdered the victim), we can ask ourselves what the world would have to be like for that to have been true. Then we move onto the final step:

4. Test those implications.

> ✔ Example [Math Processing Error]
>
> We can search the murder scene, try to find a murder weapon, run DNA analysis on the organic matter left at the scene, question the spouse about their alibi and possible motives, check their bank accounts, talk to friends and neighbors, etc. Once we have a hypothesis, in other words, that hypothesis drives the search for new evidence—it tells us what might be relevant and what irrelevant and therefore what is worth our time and what is not.

The Logic of Hypothetical Reasoning

If the spouse did it, then they must have a weak alibi. Their alibi is only verifiable by one person: the victim. So they do have a weak alibi. Therefore...they did it? Not quite.

Just because they have a weak alibi doesn't mean they did it. If that were true, anyone with a weak alibi would be guilty for everything bad that happened when they weren't busy with a verifiable activity.

Similarly, if your car's battery is dead, then it won't start. This doesn't mean that whenever your car doesn't start, the battery is dead. That would be a wild and bananas claim to make (and obviously false), but the original conditional (the first sentence in this paragraph) isn't wild and bananas. In fact, it's a pretty normal claim to make and it seems obviously true.

Let's talk briefly about the logic of hypothetical reasoning so we can discover an important truth.

If the spouse did it, then their alibi will be weak

Their alibi is weak

So, the spouse did it

This is bad reasoning. How do we know? Well, here's the logical form:

If A, then B

B

Therefore, A

This argument structure—called "affirming the consequent"—is invalid because there are countless instances of this general structure that have true premises and a false conclusion. Consider the following examples:

> ✔ Example [Math Processing Error]
>
> *If I cook, I eat well*
>
> *I ate well tonight, so I cooked.*

Example [Math Processing Error]

If Eric runs for student president, he'll become more popular.

Eric did become more popular, so he must've run for student president.

Maybe I ate well because I'm at the finest restaurant in town. Maybe I ate well because my brother cooked for me. Any of these things is possible, which is the root problem with this argument structure. It infers that one of the many possible antecedents to the conditional is the true antecedent without giving any reason for choosing or preferring this antecedent.

More concretely, affirming the consequent is the structure of an argument that states that a) one thing will explain an event, and b) that the event in question in fact occurred, and then concludes that c) the one thing that would've explained the event is the correct explanation of the event.

More concretely still, here's yet another example of affirming the consequent:

Example [Math Processing Error]

My being rich would explain my being popular

I am in fact popular,

Therefore I am in fact rich

I might be popular without having a penny to my name. People sometimes root for underdogs, or respond to the right kind of personality regardless of their socioeconomic standing, or respect a good sense of humor or athletic prowess.

If I were rich, though, that would be one potential explanation for my being popular. Rich people have nice clothes, cool cars, nice houses, and get to have the kinds of experiences that make someone a potentially popular person because everyone wants to hear the cool stories or be associated with the exciting life they lead. Perhaps, people often seem to think, *they'll get to participate in the next adventure* if they cozy up to the rich people. Rich kids in high school can also throw the best parties (if we're honest, and that's a great source of popularity).

But If I'm not rich, that doesn't mean I'm not popular. It only means that I'm not popular *because I'm rich*.

Okay, so we've established that hypothetical reasoning has the logical structure of affirming the consequent. We've further established that affirming the consequent is an *invalid* deductive argumentative structure. Where does this leave us? Is the hypothetical method *bad reasoning*?!?!?!? Nope! Luckily not all reasoning is deductive reasoning.

Remember that we're discussing *inductive* reasoning in this chapter. Inductive reasoning doesn't obey the rules of deductive logic. So it's no crime for a method of inductive reasoning to be deductively invalid. The crime against logic would be to claim that we have *certain knowledge* when we only use inductive reasoning to justify that knowledge. The upshot? Science doesn't produce certain knowledge—it produces justified knowledge, knowledge to a more or less high degree of certitude, knowledge that we can rely on and build bridges on, knowledge that almost certainly won't let us down (but it doesn't produce certain knowledge).

We can, though, with deductive certainty, *falsify* a hypothesis. Consider the murder case: if the spouse did it, then they'd have a weak alibi. That is, if the spouse did it, then they wouldn't have an airtight alibi because they'd have to be lying about where they were when the murder took place. If it turns out that the spouse *does* have an airtight alibi, then your hypothesis was wrong.

Let's take a look at the logic of falsification:

If the spouse did it, then they won't have an airtight alibi

They have an airtight alibi

So the spouse didn't do it

Now it's possible that the conditional premise (the first premise) isn't true, but we'll assume it's true for the sake of the illustration. The hypothesis was that the spouse did it and so the spouse's alibi must have some weakness.

It's also possible that our detective work hasn't been thorough enough and so the second premise is false. These are important possibilities to keep in mind. Either way, here's the logical form (a bit cleaned up and simplified):

If A, then B

Not B

Therefore not A

This is what argument pattern? That's right! You're so smart! It's *modus tollens* or "the method of denying". It's a type of argument where you deny the implications of something and thereby deny that very thing. It's a deductively valid argument form (remember from our unit on natural deduction?), so we can falsify hypotheses with deductive certainty: if your hypothesis implies something with necessity, and that something doesn't come to pass, then your hypothesis is wrong.

Your hypothesis is wrong. That is, your hypothesis *as it stands* was wrong. You might be like one of those rogue and dogged detectives in the television shows that never gives up on a hunch and ultimately discovers the truth through sheer stubbornness and determination. You might think that the spouse did it, even though they've got an airtight alibi. In that case, you'll have to **alter your hypothesis** a bit.

The process of altering a hypothesis to react to potentially falsifying evidence typically involves *adding extra hypotheses* onto your original hypothesis such that the original hypothesis no longer has the troubling implications which turned out not to be true. These extra hypotheses are called ***ad hoc* hypotheses.**

As an example, Newton's theory of gravity had one problem: it made a sort of wacky prediction. So the idea was that gravity was an instantaneous attractive force exerted by all massive bodies on all other bodies. That is, all bodies attract all other bodies regardless of distance or time. The result of this should be that all massive bodies should smack into each other over time (after all, they still have to travel towards one another). But we don't witness this. We should see things crashing towards the center of gravity of the universe at incredible speeds, but that's not what's happening. So, by the logic of falsification, Newton's theory is simply false.

But Newton had a trick up his sleeve: he claimed that God arranged things such that the heavenly bodies are so far apart from one another that they are prevented from crashing into one another. Problem solved! God put things in the right spatial orientation such that the theory of gravity is saved: they won't crash into each other because they're so far apart! Newton employed an ad hoc hypothesis to save his theory from falsification.

Abductive Reasoning

There's one more thing to discuss while we're still on the topic of hypothetical reasoning or reasoning using hypotheses. 'Abduction' is a fancy word for a process or method sometimes called "inference to the best explanation. The basic idea is that we have a bunch of evidence, we try to explain it, and we find that we could explain it in multiple ways. Then we find the "best" explanation or hypothesis and infer that this is the true explanation.

For example, say we're playing a game that's sort of like the picnic game from before. I give you a series of numbers, and then you give me more series of numbers so that I can confirm or deny that each meets the rule I have in mind. So I say:

20, 30, 40

And then you offer the following series (serieses?):

2, 3, 4

12, 22, 32

60, 90, 120

Each of these series tests a particular hypothesis. The first tests whether the important thing is that the numbers start with 2, 3, and 4. The second tests whether the rule is to add 10 each successive number in the series. The third tests a more complicated hypothesis: add half of the first number to itself to get the second number, then add one third of the second number to itself to get the third number.

Now let's say I tell you that only the third series is acceptable. What now?

Well, our hypothesis was pretty complex, but it seems pretty good. I can infer that this is the correct rule. Alternatively, I might look at other hypotheses which fit the evidence equally well: 1x, 1.5x, 2x? or maybe it's 2x, 3x, 4x? What about x, 1.5x, x*[Math Processing Error]*? These all make sense of the data, but are they equal apart from that?

Let's suppose we can't easily get more data with which to test our various hypotheses. We've got 4 to choose from and nothing in the evidence suggests that one of the hypotheses is better than the others—they all fit the evidence perfectly. What do we do?

One thing we could do is choose which hypothesis is best for reasons other than fit with the evidence. Maybe we want a simpler hypothesis, or maybe we want a more elegant hypothesis, or one which suggests more routes for investigation. These are what we might call "theoretical virtues"—they're the things we want to see in a theory. The process of abduction is the process of selecting the hypothesis that has the most to offer in terms of theoretical virtues: the simplest, most elegant, most fruitful, most general, and so on.

In science in particular, we value a few theoretical virtues over others: support by the empirical evidence available, replicability of the results in a controlled setting by other scientists, ideally mathematical precision or at least a lack of vagueness, and parsimony or simplicity in terms of the sorts of things the hypothesis requires us to believe in.

Confirmation Bias

This is a great opportunity to discuss confirmation bias, or the natural tendency we have to seek out evidence which supports our beliefs and to ignore evidence which gets in the way of our beliefs. We'll discuss cognitive biases more in Chapter 10, but since we're dealing with the relationship between evidence and belief, this seems like a good spot to pause and reflect on how our minds work.

The way our minds work naturally, it seems, is to settle on a belief and then work hard to maintain that belief whatever happens. We come to believe that global warming is anthropogenic—is caused by human activities—and then we're happy to accept a wide variety of evidence for the claim. If the evidence supports our belief, in other words, we don't take the time or energy to really investigate exactly how convincing that evidence is. If we already believe the conclusion of an inference, in other words, we are much less likely to test or analyze the inference.

Alternatively, when we see pieces of evidence or arguments that appear to point to the contrary, we are either more skeptical of that evidence or more critical of that argument. For instance, if someone notes that the Earth goes through normal cycles of warming and ice ages and warming again, we immediately will look for ways to explain how this warming period is different than others in the past. Or we might look at the period of the cycles to find out if this is happening at the "right" time in our geological history for it not to be caused by humankind. In other words, we're more skeptical of arguments or evidence that would defeat or undermine our beliefs, but we're less skeptical and critical of arguments and evidence that supports our beliefs.

Here are some questions to reflect on as you try to decide how guilty you are of confirmation bias in your own reasoning:

Questions for Reflection:

1. Which news sources do you trust? Why?
2. What's your process for exploring a topic—say a political or scientific or news topic?
3. How do you decide what to believe about a new subject?
4. When other people express an opinion about someone you don't know, do you withhold judgment? How well do you do so?
5. Are you harder on arguments and evidence that would shake up your beliefs?

9.2: Causal Reasoning

We make causal judgments all the time. We think that the world is full of things causing other things to happen. He *made* me do it. The woman lost the race *because* she was anemic. The frog jumped off the leaf, *causing* it to shake and shower dew drops onto the ground below. You're only cheating on this test because *if* you fail it, *then* you'll have to drop out of college (don't cheat, students, even if you are failing a course :)). Roasting meat until it is charred and black causes cancer.

How, though, do we know that there's a causal connection between two factors or events or things? We have to go through a particular process to demonstrate it. It's sort of a more specific version of the hypothetical method we discussed above.

First, we establish that there is in fact a *correlation* between two things. There's a correlation between alcohol use and liver disease. As one drinks more alcohol more regularly, one's chances of developing liver disease increase proportionately.

Correlation, though, is cheap. There is a hilarious website called Spurious Correlations with a collection of graphs demonstrating the odd things that are correlated with one another. Check it out! It turns out Nick Cage films correlate quite closely with deaths by drowning in a pool. Per capita mozzarella consumption correlates with civil engineering doctorate awards. So we need to get beyond mere correlation in order to posit a *causal* connection. We need to give an account and run some tests.

Second, we need to develop a causal story or an account of how they actually might be correlated. The more doctorate awards are presented, the more wine and cheese receptions there are and the more wine and cheese receptions there are, the more mozzarella cheese is consumed. Plausible? Meh, not really. But that would be how we *would* fill in the gap between the two factors.

Gamma radiation exposure causes cancer. We know it is highly correlated, but beyond that, we actually have a story of *how* it causes cancer. Gamma radiation mutates cellular DNA, which either activates or creates genes which code for uncontrolled growth. That's more or less the story of how gamma radiation and cancer are causally linked. We need this story because we need to understand how it possibly could be that gamma radiation has a causal effect on cancer.

Finally, we need to do perhaps the most important step: we need to *test* our hypothesis. We need to establish that *the more radiation one is exposed to, the more likely one is to have cancer*. Or *if there were no civil engineering doctoral degrees awarded, then no one would consume mozzarella that year*. We need to actually manipulate the variables involved until we have satisfied ourselves that by manipulating the cause we can manipulate the effect. For example, by removing the gamma exposure, we should dramatically decrease someone's chances of developing cancer. If violent video games cause violent behavior, then we should see a drop in violent behavior after we get rid of violent video games.

Necessary for establishing a causal link between two factors:

1. Establish a *prima facie* correlation
2. Develop a causal theory or story or account
3. Test for concomitant variation

If we've only established a correlation between two variables or factors and then came up with a story about how they might share a causal connection, then we've stopped short of truly knowing that there is a causal connection between the two. We have to *intervene* between the two factors to be sure that one truly causes the other. Otherwise the cause might be something that we haven't yet thought of.

9.3: Statistical Generalization

As the saying goes, there are lies, damn lies, and statistics. We won't go too far down the rabbit hole on this topic since one could teach a whole class on the logic and mathematics of statistical reasoning. All I'll do here is provide a simple account of what statistical reasoning is and then highlight a few common errors in statistical reasoning.

When generalizing using statistical methods, we always have some finite set of data and we're trying to get to a claim about the whole population. If you randomly sample one million human beings, you're probably going to end up with roughly 50/50 men and women, with non-binary folks making up a fraction as well. You might, if you think your sample is a good one, conclude that humans in general are something short of 50% likely to be men.

What makes a good sample? Two things: it must be random and it must be representative. If you want to know the attitudes of Americans about abortion rights, then sampling in Alabama isn't going to tell you much. It'll just tell you how Alabamians feel about abortion rights. That sample would be neither representative nor random. A sample being random means that the way individuals were put into the sample was done using some random method that wasn't biased in favor of any particular sub-group. If I randomly select a state on the Easter Sea Board and then use a random number generator to select a name in the phone book, then I will have randomly selected a sample, but that sample won't represent the whole United States because it will necessarily consist of people from the East Coast.

If I specifically choose people I can think of who represent every group of Americans I think are politically relevant, then I may get a perfectly representative sample, but it will be biased towards people I can think of and so are either people I know or are famous. The selection won't be random.

So we need a way of selecting members of our sample that is random, but that also ensures some measure of representativeness. Of course, if you choose 2 people, it can't be representative of the whole US. Even if you choose 1,000, you're likely to get a biased sample even if you develop a perfectly randomized selection process. So large samples are often the antidote to the possibility that even when choosing randomly, it is still possible that one's sample won't be representative.

How can statistical generalization go wrong? Here's one way: the way that data is collected can bias the *outcome*. So even if you have a perfectly random and representative sample, if you then go on to ask them whether they are in favor of "the liberal money grab", then you will likely get a result that is biased against the policy in question. Even if you are super careful, there's always the possibility of framing a question in a way that biases your subjects in favor of one answer or another.

A push poll, for instance, presents one position and sometimes even an argument in favor of that position and then asks subjects whether they agree or disagree. This is no good because it biases subjects in favor of agreeing—after all, you just gave them an argument for why they should agree.

Poll questions can also be loaded. Do you agree that we should support our troops and the wars they are fighting overseas? Even a pacifist might respond "yes" because they have nothing against the troops, they just don't like the wars. How much do you give to charity each year? Even if the true answer is "zero", few people want to admit this to a pollster. Do you agree with the unjust detainment of unaccompanied minors at the border? Who would say yes? It says "unjust" right there in the question.

Finally, one should always be wary of statistics simply because they can be manipulated so easily. Choosing one pair of factors to compare can deliver one result, while choosing a different pair will deliver a different result. You can imagine the difference if we compared lethal encounters between police and black people vs. lethal encounters between police and young black men. Or if we compared abortion rates among teenagers vs abortion rates among impoverished teenagers. Or if we compared prayer in private schools vs. prayer in schools period. Or if we compared gun violence in homes in Bel Air vs. gun violence in homes in the greater Los Angeles area. Choosing the categories to compare will bias the results in one direction or another.

Selection Bias

One thing that all too often gets in the way of good causal reasoning and good statistical generalization is a phenomenon called "selection bias."

Do married men really live longer? Actually, yeah, it turns out that they do. Is this because marriage *causes* their longevity or is it because the type of man who gets married is more likely to be the type of man who would have lived longer anyways?

Let's break it down. There are two possibilities. When you read the headline "Married Men Live Longer, Study says" you are likely to think that it is saying something like "marriage makes men live longer through the effect that having a spouse and/or children has on one's likelihood to engage in risky activities." Something like that. There's a mechanism that makes men tend to live longer when they are married, and that mechanism is part of marriage itself: the responsibility that comes with being married, or the fact that someone else is looking out for your health, or maybe there's a gender dynamic in heterosexual marriages in particular that involves a connection between living with a woman and living longer. Each of these possibilities suggests new routes of exploration and new kinds of evidence we would want to have.

But that's not the only possibility. Here's another one: there are only certain men who are likely to get married in the first place. The die-hard party machine who will never settle down is also likely to die young of liver disease. The men who die tragic deaths in their early 20's won't get married either. The men who never quite get their lives together enough to feel like they could be in a long term committed relationship might also be more likely to die of heart disease due to a sedentary lifestyle. These are mostly just suppositions. What is important here is recognizing the *possibility* that being a man who gets married may *already* put you in the category of men who will live longer *whether they get married or not*.

Selection bias happens when the sample we generalize from isn't representative of the total population in some important respect. Some factor makes it so that we aren't generalizing from a truly random sample or makes it so that what appears to be a causal relationship is instead just a selection relationship (like marriage *selects* rather than *causes* long-lived men).

Suppose you are trying to do a national survey of political opinions. Here's a method: choose addresses at random and then go to those addresses in the middle of the day and talk to whoever opens the door. Any problems you see with this strategy?

Well you may get some stay-home parents, or retirees. You sure won't get many single parents or homes with two working parents. Do you see how your selection strategy isn't truly random? Sure, you chose addresses at random, but then you ignored anyone who wasn't home when you showed up. That selects for certain people and so you'll have a biased sample.

Self-selection is also a form of selection bias. Fox News polls will deliver overwhelmingly conservative results. Yelp reviews are going to be biased in favor of loyal customers and angry customers. Those in the middle aren't likely to post reviews. American Idol is not a good poll of what Americans think because the only people who call in are those who already watch the show and care enough to make an attempt to vote.

When you hear a causal claim or statistical claim like "people who x are more likely to y" on the news, or read it online or in the paper, it is always important to ask yourself: could this be selection bias? Are women more likely to get osteoporosis or are the people who get tested for osteoporosis already more likely to be women? For instance, women have a higher life expectancy, so perhaps there are more octogenarian women. I think they've accounted for this in their studies already (or perhaps these claims are simply false :), what do I know?), but a dose of skepticism is healthy.

Understanding selection bias is a tool to put in your toolkit. Developing the habit of checking for the possibility of selection bias makes you a better thinker.

Selective Reporting

Finally, we should discuss the most common tactic used in the political sphere: reporting the same data in different ways to achieve different rhetorical goals. You might say "feeding the homeless in every state in the union would cost 50 million a year." Or you might say "feeding the homeless in every state in the union would cost .0001% of the Pentagon's yearly budget." When we talk in comparative terms, often we have a deeper grasp of what these large numbers mean.

Examples:

Example [Math Processing Error]

90% of people who take our medicine recover from their colds in under a week

Vs. 90% of people recover from their colds in under a week

Example [Math Processing Error]

Only .05% of Russian Immigrants voted in the last election

Vs. 98% of Russian immigrants who have US citizenship voted in the last election

Example [Math Processing Error]

Selected test subjects showed a 200% increase in efficacy

Vs. 2 test subjects out of the 10,000 tested showed a 200% increase in efficacy

---

9.4: Analogical Reasoning

It may be the most common type of inductive reasoning. It's sort of built into our cognitive systems. What is it? Reasoning *by analogy*. We use the known or observed similarities of things to posit further similarities. In this edition of the textbook, we'll cover this quite briefly. In future editions, I hope to expand this section to a more complete analysis.

The first thing to do is to look at some examples of analogical reasoning or reasoning by analogy:

Example [Math Processing Error]

If you increase the gauge of copper wire, you increase the maximum amperage that wire can conduct. Therefore, since water and electricity share many commonalities particularly with respect to their flow, we should see an increase in flow rate with an increase in pipe diameter.

Example [Math Processing Error]

Electrons and Photons are both point particles, therefore we should expect electrons to travel at or near the speed of light.

Example [Math Processing Error]

My boyfriend is acting just like that character on television that is cheating on their partner, so my boyfriend is probably cheating.

Example [Math Processing Error]

Samsung produces both the Galaxy S8 and the Galaxy S9, so I would expect their basic operating systems to be quite similar.

Example [Math Processing Error]

Northern California and Northwestern Italy have very similar climates, so probably olives and grapes will grow well in Northern California.

One can see how this isn't perfect reasoning, right? There are going to be loads of counterexamples to each of these types of analogical reasoning. But some of these particular examples turn out to have true conclusions and they seem to have identified a commonality between the two **analogues** (the things being compared in the analogy) that in fact *makes the conclusion true*. For instance, it *is* the similar climates of Northern California and Tuscany that makes for quite similar growing conditions and therefore the success of quite similar crops. It *is* the properties shared by all point particles that leads them to travel at or near the speed of light (it's not impossible I'm wrong about this, I'm no physicist). It *is* the common manufacturer that makes Samsung phones similar. In other words: the commonalities between things often do cause them to have further commonalities.

Arguments by analogy have a certain structure in common. It's worth investigating this a bit further:

Analogues	
Attribute(s) in Common	
Attribute(s) inferred	
Link between attributes	

Let's take an example from above and analyze it according to this schema:

Example [Math Processing Error]

If you increase the gauge of copper wire, you increase the maximum amperage that wire can conduct. Therefore, since water and electricity share many commonalities particularly with respect to their flow, we should see an increase in flow rate with an increase in pipe diameter.

Now we identify the analogues first:

Analogues	Electricity flow vs. Water flow

It can be tricky to get the analogues just right, but the Attribute in Common can help us narrow it down.

Attribute(s) in Common	Characteristics of flow like amperage/flow rate; voltage/pressure; and resistance/friction.

Next we identify what is being inferred about the pair of analogues: what else are they supposed to have in common on the basis of their known similarities?

Attribute(s) inferred	Increase in diameter of conduit equals increase in amperage/flow rate.

Finally, we might wonder what is it about the Attributes in Common that might lead us to believe that the analogues will share the further attribute(s)? What is the causal link? Here's my best guess about what's going on here (again, I'm no physicist):

Link between attributes	The flow of electrons in fact shares many properties with the flow of water molecules: both pipes and copper wires allow a certain diameter of flowing medium (electrons or water molecules) to flow through one particular slice at a given time. Only a certain amount of electrons and a certain amount of water molecules can fit through a hose/wire of a given length.

This is a tricky example, since the right answer relies on physics I am not an expert in (not by a long stretch). What matters at this point is that we're able to identify good and bad analogies by means of these tools of analysis. I won't expect to find a link between attributes forthcoming in cases of bad analogical reasoning. Here's an example:

Example [Math Processing Error]

Both of my pairs of shoes are blue, so I'd expect one [Stiletto heels] to perform just as well on a hike as the other [Chaco Sandals].

Let's break down this analogy using our handy chart:

Analogues	Stiletto Heels vs. Chaco Sandals
Attribute(s) in Common	Blue, they are footwear
Attribute(s) inferred	Performance on a hike
Link between attributes	They are at least footwear, so they will perform better than most nonfootwear. Better than a slice of bread, for instance. But that's it.

There is no link between color and function when it comes to shoes and so the analogy is a bad one. Even if the link is unknown or a bit mysterious (again, I'm no physicist, so the chart about electrons is probably a bit off), we can usually still identify *something like* the right link between attributes if we understand anything about how the analogy works. In the case of a bad analogy, we can make up wild stories, but at the end of the day we'll most likely realize that there simply is no link between the attributes in question.

What about when the attribute in common isn't connected to the attribute inferred? Well, it's probably not a good bit of analogical reasoning, even if it ends up having a true conclusion. Here's an example of this sort of thing:

Example [Math Processing Error]

Rolexes and Teslas are both expensive. Teslas are very durable, so it follows that Rolexes are very durable.

Let's break it down:

Analogues	Rolex watches vs. Tesla cars
Attribute(s) in Common	They are expensive consumer goods
Attribute(s) inferred	Durability
Link between attributes	I guess when things are expensive the manufacturer can spend more money on durable materials and the like, but we all know there are durable cheap goods and fragile expensive goods, so we have to say more than just their price point to infer durability.

Sure, they are both durable, but no, they are not both durable *because they are expensive* nor vice versa. They are expensive because of market forces like supply and demand, and they are durable because they are well-made. The two don't really share much of a link, even if there is some sort of a weak link between the two.

In closing, analogical reasoning starts from some commonalities between two or more sets of things (analogues), and then it proceeds by inferring further commonalities. It is successful when either a) the common attributes cause or explain the inferred attributes; or b) the common and inferred attributes are both caused or explained by something else about the two analogues. It is unsuccessful when neither (a) nor (b) is true. Using the chart on offer in this section can be helpful in breaking down analogical reasoning and identifying just what is being claimed. Sometimes analyzing it like this can lay bare the faults in such forms of reasoning.

The following is a good discussion of the evaluation of analogical arguments from Matthew Knachels' Fundamental Methods of Logic: (CC-BY-4.0 License):

The Evaluation of Analogical Arguments

Unlike in the case of deduction, we will not have to learn special techniques to use when evaluating these sorts of arguments. It's something we already know how to do, something we typically do automatically and unreflectively. The purpose of this section, then, is not to learn a new skill, but rather subject a practice we already know how to engage in to critical scrutiny. We evaluate analogical arguments all the time without thinking about how we do it. We want to achieve a metacognitive perspective on the practice of evaluating arguments from analogy; we want to think about a type of thinking that we typically engage in without much conscious deliberation. We want to identify the criteria that we rely on to evaluate analogical reasoning—criteria that we apply without necessarily realizing that we're applying them. Achieving such metacognitive awareness is useful insofar as it makes us more self-aware, critical, and therefore effective reasoners.

Analogical arguments are inductive arguments. They give us reasons that are supposed to make their conclusions more probable. How probable, exactly? That's very hard to say. How probable was it that I would like The Wolf of Wall Street given that I had liked the other four Scorsese/DiCaprio collaborations? I don't know. How probable is it that it's wrong to eat pork given that it's wrong to eat dogs and dolphins? I really don't know. It's hard to imagine how you would even begin to answer that question.

As we mentioned, while it's often impossible to evaluate inductive arguments by giving a precise probability of its conclusion, it is possible to make relative judgments about strength and weakness. Recall, new information can change the probability of the conclusion of an inductive argument. We can make relative judgments of like this: if we add this new information as a premise, the new argument is stronger/weaker than the old argument; that is, the new information makes the conclusion more/less likely.

It is these types of relative judgments that we make when we evaluate analogical reasoning. We compare different arguments —with the difference being new information in the form of an added premise, or a different conclusion supported by the same premises—and judge one to be stronger or weaker than the other. Subjecting this practice to critical scrutiny, we can identify six criteria that we use to make such judgments.

We're going to be making relative judgments, so we need a baseline argument against which to compare others. Here is such an argument:

Alice has taken four Philosophy courses during her time in college. She got an A in all four. She has signed up to take another Philosophy course this semester. I predict she will get an A in that course, too.

This is a simple argument from analogy, in which the future is predicted based on past experience. It fits the schema for analogical arguments: the new course she has signed up for is designated by ‘c’; the property we’re predicting it has (Q) is that it is a course Alice will get an A in; the analogues are the four previous courses she’s taken; what they have in common with the new course (P1) is that they are also Philosophy classes; and they all have the property Q—Sally got an A in each.

Anyway, how strong is the baseline argument? How probable is its conclusion in light of its premises? I have no idea. It doesn’t matter. We’re now going to consider tweaks to the argument, and the effect that those will have on the probability of the conclusion. That is, we’re going to consider slightly different arguments, with new information added to the original premises or changes to the prediction based on them, and ask whether these altered new arguments are stronger or weaker than the baseline argument. This will reveal the six criteria that we use to make such judgments. We’ll consider one criterion at a time.

Number of Analogues

Suppose we alter the original argument by changing the number of prior Philosophy courses Alice had taken. Instead of Alice having taken four philosophy courses before, we’ll now suppose she has taken 14. We’ll keep everything else about the argument the same: she got an A in all of them, and we’re predicting she’ll get an A in the new one. Are we more or less confident in the conclusion—the prediction of an A—with the altered premise? Is this new argument stronger or weaker than the baseline argument?

It’s stronger! We’ve got Alice getting an A 14 times in a row instead of only four. That clearly makes the conclusion more probable. (How much more? Again, it doesn’t matter.)

What we did in this case is add more analogues. This reveals a general rule: other things being equal, the more analogues in an analogical argument, the stronger the argument (and conversely, the fewer analogues, the weaker). The number of analogues is one of the criteria we use to evaluate arguments from analogy.

Variety of Analogues

You’ll notice that the original argument doesn’t give us much information about the four courses Alice succeeded in previously and the new course she’s about to take. All we know is that they’re all Philosophy courses. Suppose we tweak things. We’re still in the dark about the new course Alice is about to take, but we know a bit more about the other four: one was a course in Ancient Greek Philosophy; one was a course on Contemporary Ethical Theories; one was a course in Formal Logic; and the last one was a course in the Philosophy of Mind. Given this new information, are we more or less confident that she will succeed in the new course, whose topic is unknown to us? Is the argument stronger or weaker than the baseline argument?

It is stronger. We don’t know what kind of Philosophy course Alice is about to take, but this new information gives us an indication that it doesn’t really matter. She was able to succeed in a wide variety of courses, from Mind to Logic, from Ancient Greek to Contemporary Ethics. This is evidence that Alice is good at Philosophy generally, so that no matter what kind of course she’s about to take, she’ll probably do well in it.

Again, this points to a general principle about how we evaluate analogical arguments: other things being equal, the more variety there is among the analogues, the stronger the argument (and conversely, the less variety, the weaker).

Number of Similarities

In the baseline argument, the only thing the four previous courses and the new course have in common is that they’re Philosophy classes. Suppose we change that. Our newly tweaked argument predicts that Alice will get an A in the new course, which, like the four she succeeded in before, is cross-listed in the Department of Religious Studies and covers topics in the Philosophy of Religion. Given this new information—that the new course and the four older courses were similar in ways we weren’t aware of—are we more or less confident in the prediction that Alice will get another A? Is the argument stronger or weaker than the baseline argument?

Again, it is stronger. Unlike the last example, this tweak gives us new information both about the four previous courses and the new one. The upshot of that information is that they’re more similar than we knew; that is, they have more properties in common. To P1 = ‘is a Philosophy course’ we can add P2 = ‘is cross-listed with Religious Studies’ and P3 = ‘covers topics in Philosophy of Religion’. The more properties things have in common, the stronger the analogy between them. The stronger the analogy, the stronger the argument based on that analogy. We now know not just that Alice did well in not just in Philosophy classes—but specifically in classes covering the Philosophy of Religion; and we know that the new class she’s taking is also a

Philosophy of Religion class. I'm much more confident predicting she'll do well again than I was when all I knew was that all the classes were Philosophy; the new one could've been in a different topic that she wouldn't have liked.

General principle: other things being equal, the more properties involved in the analogy—the more similarities between the item in the conclusion and the analogues—the stronger the argument (and conversely, the fewer properties, the weaker).

Number of Differences

An argument from analogy is built on the foundation of the similarities between the analogues and the item in the conclusion—the analogy. Anything that weakens that foundation weakens the argument. So, to the extent that there are differences among those items, the argument is weaker.

Suppose we add new information to our baseline argument: the four Philosophy courses Alice did well in before were all courses in the Philosophy of Mind; the new course is about the history of Ancient Greek Philosophy. Given this new information, are we more or less confident that she will succeed in the new course? Is the argument stronger or weaker than the baseline argument? Clearly, the argument is weaker. The new course is on a completely different topic than the other ones. She did well in four straight Philosophy of Mind courses, but Ancient Greek Philosophy is quite different. I'm less confident that she'll get an A than I was before.

If I add more differences, the argument gets even weaker. Supposing the four Philosophy of Mind courses were all taught by the same professor (the person in the department whose expertise is in that area), but the Ancient Greek Philosophy course is taught by someone different (the department's specialist in that topic). Different subject matter, different teachers: I'm even less optimistic about Alice's continued success.

Generally speaking, other things being equal, the more differences there are between the analogues and the item in the conclusion, the weaker the argument from analogy.

Relevance of Similarities and Differences

Not all similarities and differences are capable of strengthening or weakening an argument from analogy, however. Suppose we tweak the original argument by adding the new information that the new course and the four previous courses all have their weekly meetings in the same campus building. This is an additional property that the courses have in common, which, as we just saw, other things being equal, should strengthen the argument. But other things are not equal in this case. That's because it's very hard to imagine how the location of the classroom would have anything to do with the prediction we're making—that Alice will get an A in the course. Classroom location is simply not relevant to success in a course.[1] Therefore, this new information does not strengthen the argument. Nor does it weaken it; I'm not inclined to doubt Alice will do well in light of the information about location. It simply has no effect at all on my appraisal of her chances.

Similarly, if we tweak the original argument to add a difference between the new class and the other four, to the effect that while all of the four older classes were in the same building, while the new one is in a different building, there is no effect on our confidence in the conclusion. Again, the building in which a class meets is simply not relevant to how well someone does.

Contrast these cases with the new information that the new course and the previous four are all taught by the same professor. Now that strengthens the argument! Alice has gotten an A four times in a row from this professor—all the more reason to expect she'll receive another one. This tidbit strengthens the argument because the new similarity—the same person teaches all the courses—is relevant to the prediction we're making—that Alice will do well. Who teaches a class can make a difference to how students do—either because they're easy graders, or because they're great teachers, or because the student and the teacher are in tune with one another, etc. Even a difference between the analogues and the item in the conclusion, with the right kind of relevance, can strengthen an argument. Suppose the other four philosophy classes were taught be the same teacher, but the new one is taught by a TA—who just happens to be her boyfriend. That's a difference, but one that makes the conclusion—that Alice will do well—more probable.

Generally speaking, careful attention must be paid to the relevance of any similarities and differences to the property in the conclusion; the effect on strength varies.

Modesty/Ambition of the Conclusion

Suppose we leave everything about the premises in the original baseline argument the same: four Philosophy classes, an A in each, new Philosophy class. Instead of adding to that part of the argument, we'll tweak the conclusion. Instead of predicting that Alice will get an A in the class, we'll predict that she'll pass the course. Are we more or less confident that this prediction will come true? Is the new, tweaked argument stronger or weaker than the baseline argument?

It's stronger. We are more confident in the prediction that Alice will pass than we are in the prediction that she will get another A, for the simple reason that it's much easier to pass than it is to get an A. That is, the prediction of passing is a much more modest prediction than the prediction of an A.

Suppose we tweak the conclusion in the opposite direction—not more modest, but more ambitious. Alice has gotten an A in four straight Philosophy classes, she's about to take another one, and I predict that she will do so well that her professor will suggest that she publish her term paper in one of the most prestigious philosophical journals and that she will be offered a three-year research fellowship at the Institute for Advanced Study at Princeton University. That's a bold prediction! Meaning, of course, that it's very unlikely to happen. Getting an A is one thing; getting an invitation to be a visiting scholar at one of the most prestigious academic institutions in the world is quite another. The argument with this ambitious conclusion is weaker than the baseline argument.

General principle: the more modest the argument's conclusion, the stronger the argument; the more ambitious, the weaker.

Refutation by Analogy

We can use arguments from analogy for a specific logical task: refuting someone else's argument, showing that it's bad. Recall the case of deductive arguments. To refute those—to show that they are bad, i.e., invalid—we had to produce a counterexample—a new argument with the same logical form as the original that was obviously invalid, in that its premises were in fact true and its conclusion in fact false. We can use a similar procedure to refute inductive arguments. Of course, the standard of evaluation is different for induction: we don't judge them according to the black and white standard of validity. And as a result, our judgments have less to do with form than with content. Nevertheless, refutation along similar lines is possible, and analogies are the key to the technique.

To refute an inductive argument, we produce a new argument that's obviously bad—just as we did in the case of deduction. We don't have a precise notion of logical form for inductive arguments, so we can't demand that the refuting argument have the same form as the original; rather, we want the new argument to have an analogous form to the original. The stronger the analogy between the refuting and refuted arguments, the more decisive the refutation. We cannot produce the kind of knock-down refutations that were possible in the case of deductive arguments, where the standard of evaluation—validity—does not admit of degrees of goodness or badness, but the technique can be quite effective.

Consider the following:

> "Duck Dynasty" star and Duck Commander CEO Willie Robertson said he supports Trump because both of them have been successful businessmen and stars of reality TV shows.
>
> By that logic, does that mean Hugh Hefner's success with "Playboy" and his occasional appearances on "Bad Girls Club" warrant him as a worthy president? Actually, I'd still be more likely to vote for Hefner than Trump.[2]

The author is refuting the argument of Willie Robertson, the "Duck Dynasty" star. Robertson's argument is something like this: Trump is a successful businessman and reality TV star; therefore, he would be a good president. To refute this, the author produces an analogous argument—Hugh Hefner is a successful businessman and reality TV star; therefore, Hugh Hefner would make a good president—that he regards as obviously bad. What makes it obviously bad is that it has a conclusion that nobody would agree with: Hugh Hefner would make a good president. That's how these refutations work. They attempt to demonstrate that the original argument is lousy by showing that you can use the same or very similar reasoning to arrive at an absurd conclusion.

Here's another example, from a group called "Iowans for Public Education". Next to a picture of an apparently well-to-do lady is the following text:

> "My husband and I have decided the local parks just aren't good enough for our kids. We'd rather use the country club, and we are hoping state tax dollars will pay for it. We are advocating for Park Savings Accounts, or PSAs. We promise to no longer use the local parks. To hell with anyone else or the community as a whole. We want our tax dollars to be used to make the best choice for our family."

> Sound ridiculous? Tell your legislator to vote NO on Education Savings Accounts (ESAs), aka school vouchers.

The argument that Iowans for Public Education put in the mouth of the lady on the poster is meant to refute reasoning used by advocates for “school choice”, who say that they ought to have the right to opt out of public education and keep the tax dollars they would otherwise pay for public schools and use it to pay to send their kids to private schools. A similar line of reasoning sounds pretty crazy when you replace public schools with public parks and private schools with country clubs.

Since these sorts of refutations rely on analogies, they are only as strong as the analogy between the refuting and refuted arguments. There is room for dispute on that question. Advocates for school vouchers might point out that schools and parks are completely different things, that schools are much more important to the future prospects of children, and that given the importance of education, families should have to right choose what they think is best. Or something like that. The point is, the kinds of knock-down refutations that were possible for deductive arguments are not possible for inductive arguments. There is always room for further debate.

---

[1] I’m sure someone could come up with some elaborate backstory for Alice according to which the location of the class somehow makes it more likely that she will do well, but set that aside. No such story is on the table here.

[2] Austin Faulds, “Weird celebrity endorsements fit for weird election,” Indiana Daily Student, 10/12/16, http://www.idsnews.com/article/2016/...weird-election.

9.5: Fallacies of Induction

The fallacies of induction are all failures in reasoning about the messy world of cause and effect, contingent facts of the universe, and generalizations about kinds of things in the world. In each case, an argument is put forth using evidence incorrectly, or making bad predictions, or generalizing improperly.

Appeal to Ignorance

We can't prove either way whether there is or isn't an all-powerful god, you might suppose. (Intelligent people disagree). If that's true, then it seems like we're free to believe in an all-powerful loving god, right? Well... in a sense. You're free to believe whatever you like, but we shouldn't pretend that you're *justified* in believing in a god on the basis of the lack of proof against the existence of god. You could cite the exact same evidence (i.e. the lack of conclusive evidence for either side) to justify believing in no god. Or, as is sometimes done, to justify believing in a Flying Spaghetti Monster that created the universe.

Whether or not you believe in a creator god, therefore, cannot depend on the seeming fact that no one has proved that there isn't a god. There might be other reasons for believing in a god, but this ain't one of them.

This example illustrates the structure of the argument from ignorance. Essentially, the basic argument pattern looks like this:

We don't know whether proposition x is true or false

Therefore, it's true

or

Therefore, it's false

This is a bad argument pattern because the fact that we don't know the truth of the matter is reason for *withholding judgment* and not coming to hold a determinate belief. It's not reason for or against believing something.

Sometimes, it's okay not to have an opinion or belief about a particular subject matter.

The following is from: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

This is a particularly egregious and perverse fallacy. In essence, it's an inference from premises to the effect that there's a lack of knowledge about some topic to a definite conclusion about that topic. We don't know; therefore, we know!

Of course, put that baldly, it's plainly absurd; actual instances are more subtle. The fallacy comes in a variety of closely related forms. It will be helpful to state them in bald/absurd schematic fashion first, then elucidate with more subtle real-life examples.

The first form can be put like this:

[Math Processing Error]

That sounds silly, but consider an example: those "documentary" programs on cable TV about aliens. You know, the ones where they suggest that extraterrestrials built the pyramids or something (there are books and websites, too). How do they get you to believe that crazy theory? By creating mystery! By pointing to facts that nobody can explain. The Great Pyramid at Giza is aligned (almost) exactly with the magnetic north pole! On the day of the summer solstice, the sun sets exactly between two of the pyramids! The height of the Great Pyramid is (almost) exactly one one-millionth the distance from the Earth to the Sun! How could the ancient Egyptians have such sophisticated astronomical and geometrical knowledge? Why did the Egyptians, careful record-keepers in (most) other respects, (apparently) not keep detailed records of the construction of the pyramids? Nobody knows. Conclusion: aliens built the pyramids.

In other words, there are all sorts of (sort of) surprising facts about the pyramids, and nobody knows how to explain them. From these premises, which establish only our ignorance, we're encouraged to conclude that we know something: aliens built the pyramids. That's quite a leap—too much of a leap.

Another form this fallacy takes can be put crudely thus:

[Math Processing Error]

The word ‘prove’ is in all-caps because stressing it is the key to this fallacious argument: the standard of proof is set impossibly high, so that almost no amount of evidence would constitute a refutation of the conclusion.

An example will help. There are lots of people who claim that evolutionary biology is a lie: there’s no such thing as evolution by natural selection, and it’s especially false to claim that humans evolved from earlier species, that we share a common ancestor with apes. Rather, the story goes, the Bible is literally true: the Earth is only about 6,000 years old, and humans were created as-is by God just as the Book of Genesis describes. The Argument from Ignorance is one of the favored techniques of proponents of this view. They are especially fond of pointing to “gaps” in the fossil record—the so-called “missing link” between humans and a pre-human, ape-like species—and claim that the incompleteness of the fossil record vindicates their position.

But this argument is an instance of the fallacy. The standard of proof—a complete fossil record without any gaps—is impossibly high. Evolution has been going on for a LONG time (the Earth is actually about 4.5 billion years old, and living things have been around for at least 3.5 billion years). So many species have appeared and disappeared over time that it’s absurd to think that we could even come close to collecting fossilized remains of anything but the tiniest fraction of them. It’s hard to become a fossil, after all: a creature has to die under special circumstances to even have a chance for its remains to do anything than turn into compost. And we haven’t been searching for fossils in a systematic way for very long (only since the mid-1800s or so). It’s no surprise that there are gaps in the fossil record, then. What’s surprising, in fact, is that we have as rich a fossil record as we do. Many, many transitional species have been discovered, both between humans and their ape-like ancestors, and between other modern species and their distant forbears (whales used to be land-based creatures, for example; we know this (in part) from the fossils of early proto-whale species with longer and longer rear hip- and leg-bones).

We will never have a fossil record complete enough to satisfy skeptics of evolution. But their standard is unreasonably high, so their argument is fallacious. Sometimes they put it even more simply: nobody was around to witness evolution in action; therefore, it didn’t happen. This is patently absurd, but it follows the same pattern: an unreasonable standard of proof (witnesses to evolution in action; impossible, since it takes place over such a long period of time), followed by the leap to the unwarranted conclusion.

Yet another version of the Argument from Ignorance goes like this:

[Math Processing Error]

Of course lack of imagination on the part of an individual isn’t evidence for or against a proposition, but people often argue this way. A (hilarious) example comes from the rap duo Insane Clown Posse in their 2009 single, “Miracles”. Here’s the line:

Example *[Math Processing Error]*

Water, fire, air and dirt

F**king magnets, how do they work?

And I don’t wanna talk to a scientist

Y’all mother**kers lying, and getting me pissed.

Violent J and Shaggy 2 Dope can’t understand how there could be a scientific, non-miraculous explanation for the workings of magnets. They conclude, therefore, that magnets are miraculous.

A final form of the Argument from Ignorance can be put crudely thus:

[Math Processing Error]

You may have heard the slogan, “Absence of evidence is not evidence of absence.” This is an attempt to sum up this version of the fallacy. But it’s not quite right. What it should say is that absence of evidence is not always definitive evidence of absence. An example will help illustrate the idea. During the 2016 presidential campaign, a reporter (David Fahrentold) took to Twitter to announce that despite having “spent weeks looking for proof that [Donald Trump] really does give millions of his own [money] to charity...” he could only find one donation, to the NYC Police Athletic League. Trump has claimed to have given millions of dollars to charities over the years. Does this reporter’s failure to find evidence of such giving prove that Trump’s claims about his charitable donations are false? No. To rely only on this reporter’s testimony to draw such a conclusion would be to commit the fallacy.

However, the failure to uncover evidence of charitable giving does provide some reason to suspect Trump's claims may be false. How much of a reason depends on the reporter's methods and credibility, among other things.9 But sometimes a lack of evidence can provide strong support for a negative conclusion. This is an inductive argument; it can be weak or strong. For example, despite multiple claims over many years (centuries, if some sources can be believed), no evidence has been found that there's a sea monster living in Loch Ness in Scotland. Given the size of the body of water, and the extensiveness of the searches, this is pretty good evidence that there's no such creature—a strong inductive argument to that conclusion. To claim otherwise—that there is such a monster, despite the lack of evidence—would be to commit the version of the fallacy whereby one argues "You can't PROVE I'm wrong; therefore, I'm right," where the standard of proof is unreasonably high.

One final note on this fallacy: it's common for people to mislabel certain bad arguments as arguments from ignorance; namely, arguments made by people who obviously don't know what the heck they're talking about. People who are confused or ignorant about the subject on which they're offering an opinion are liable to make bad arguments, but the fact of their ignorance is not enough to label those arguments as instances of the fallacy. We reserve that designation for arguments that take the forms canvassed above: those that rely on ignorance—and not just that of the arguer, but of the audience as well—as a premise to support the conclusion.

Slippery Slope

Figure *[Math Processing Error]*: First he's harnessing and whipping the dog, next he'll be murdering folks! (Image Credit: Otto Speckter in *Picture Fables*)

From Matthew J. Van Cleave's Introduction to Logic and Critical Thinking, version 1.4, pp. 189-195 Creative Commons Attribution 4.0 International License.

The causal slippery slope fallacy is committed when one event is said to lead to some other (usually disastrous) event via a chain of intermediary events. If you have ever seen Direct TV's "get rid of cable" commercials, you will know exactly what I'm talking about. (If you don't know what I'm talking about you should Google it right now and find out. They're quite funny.) Here is an example of a causal slippery slope fallacy (it is adapted from one of the Direct TV commercials):

Example [Math Processing Error]

If you use cable, your cable will probably go on the fritz. If your cable is on the fritz, you will probably get frustrated. When you get frustrated you will probably hit the table. When you hit the table, your young daughter will probably imitate you. When your daughter imitates you, she will probably get thrown out of school. When she gets thrown out of school, she will probably meet undesirables. When she meets undesirables, she will probably marry undesirables. When she marries undesirables, you will probably have a grandson with a dog collar. Therefore, if you use cable, you will probably have a grandson with dog collar.

This example is silly and absurd, yes. But it illustrates the causal slippery slope fallacy. Slippery slope fallacies are always made up of a series of conjunctions of probabilistic conditional statements that link the first event to the last event. A causal slippery slope fallacy is committed when one assumes that just because each individual conditional statement is probable, the conditional that links the first event to the last event is also probable. Even if we grant that each "link" in the chain is individually probable, it doesn't follow that the whole chain (or the conditional that links the first event to the last event) is probable. Suppose, for the sake of the argument, we assign probabilities to each "link" or conditional statement, like this. (I have italicized the consequents of the conditionals and assigned high conditional probabilities to them. The high probability is for the sake of the argument; I don't actually think these things are as probable as I've assumed here.)

- If you use cable, then your cable will probably go on the fritz (.9)
- If your cable is on the fritz, then you will probably get angry (.9)
- If you get angry, then you will probably hit the table (.9)
- If you hit the table, your daughter will probably imitate you (.8)
- If your daughter imitates you, she will probably be kicked out of school (.8)
- If she is kicked out of school, she will probably meet undesirables (.9)
- If she meets undesirables, she will probably marry undesirables (.8)
- If she marries undesirables, you will probably have a grandson with a dog collar (.8)

However, even if we grant the probabilities of each link in the chain is high (80-90% probable), the conclusion doesn't even reach a probability higher than chance. Recall that in order to figure the probability of a conjunction, we must multiply the probability of each conjunct:

[Math Processing Error]

That means the probability of the conclusion (i.e., that if you use cable, you will have a grandson with a dog collar) is only 27%, despite the fact that each conditional has a relatively high probability!

Texas Sharpshooter

As the story goes, once there was a man in Texas who shot at his barn door with his rifle. When he had unloaded ten rounds, he walked up the to door, found a cluster of bullet holes that were particularly tightly clustered, and painted a bullseye around them. No one would credit him with being a sharpshooter, would they? After all, he didn't *actually* aim at a bullseye and hit it. He drew the bullseye around his shots!

This story relates to a particular way of using *evidence* to demonstrate a conclusion. Normally, we would hope that someone would take in all of the available evidence about a particular subject, weigh its relative credibility, and then come to a conclusion. The Texas Sharpshooter fallacy happens when someone already knows which conclusion they'd like to prove and then selects evidence which supports that conclusion. They've done the process *backwards.* The analogy is a little weird, but the idea is that the painting of the bullseye is selecting which evidence to take into account. If you only weigh the evidence which supports the conclusion you like (or in the story, if you only draw the target around the bullet holes that looked good) then you'd be disregarding other evidence for no other reason than that it got in the way of you concluding what you wanted to conclude.

The paradigm example of this is when you let your confidence in a particular conclusion change the way you treat evidence. Here's an example:

Example *[Math Processing Error]*

I know that Vaccines cause autism, so the multiple review articles concluding that there is no link between the two must have been bought and paid for by big pharma!

Instead of looking at the evidence and letting it determine what conclusion we draw, instead we're letting our fixed conclusion determine how we treat the evidence! It's backwards, just like Texas Sharpshooting!

This fallacy might also be called the fallacy of **Cherry-Picking Evidence** because you've selected only some evidence (you've "cherry picked").

Here's an example. Say you think vaccines are unsafe. If that's a belief you're committed to, you might only pay attention to evidence like anecdotes about apparent vaccine injuries, the medical professionals who make claims about vaccines being dangerous, and the apparent empirical evidence that connects vaccines and illnesses of various kinds. You'd likely ignore, discount, or explain away all of the evidence which seems to show that vaccines aren't connected with illness or injury in a significant way. This isn't so much a judgment on this particular belief as it is a description of what might be going on in someone's mind.

Figure [Math Processing Error]: Annie Oakley was no Texas Sharpshooter. She was a real sharpshooter.

The core problem here is *letting your desired conclusion determine which evidence you take into account or how you treat evidence.*

Any use of anecdotal evidence—a single one-off story about individuals that is supposed to be evidence for a general claim—is likely to be an instance of this fallacy, since it's usually easy to find an anecdote to support *any* claim. Anecdotes aren't evidence for general claims. At best, they're illustrations of general points.

Here's another example of using selective evidence to get the conclusion you want:

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

Quoting out of Context

Another way to obscure or alter the meaning of what someone actually said is to quote them selectively. Remarks taken out of their proper context might convey a different meaning than they did within that context.

Consider a simple example: movie ads. These often feature quotes from film critics, which are intended to convey the impression that the movie was well-liked by them. "Critics call the film 'unrelenting', 'amazing', and 'a one-of-a-kind movie experience'", the ad might say. That sounds like pretty high praise. I think I'd like to see that movie. That is, until I read the actual review from which those quotes were pulled:

> I thought I'd seen it all at the movies, but even this jaded reviewer has to admit that this film is something new, a one-of-a-kind movie experience: two straight hours of unrelenting, snooze-inducing mediocrity. I find it amazing that not one single aspect of this movie achieves even the level of "eh, I guess that was OK."

The words 'unrelenting' and 'amazing'—and the phrase 'a one-of-a-kind movie experience'—do in fact appear in that review. But situated in their original context, they're doing something completely different than the movie ad would like us to believe.

Politicians often quote each other out of context to make their opponents look bad. In the 2012 presidential campaign, both sides did it rather memorably. The Romney campaign was trying to paint President Obama as anti-business. In a campaign speech, Obama once said the following:

> If you've been successful, you didn't get there on your own. You didn't get there on your own. I'm always struck by people who think, well, it must be because I was just so smart. There are a lot of smart people out there. It must be because I worked harder than everybody else. Let me tell you something: there are a whole bunch of hardworking people out there. If you've got a business, you didn't build that. Somebody else made that happen.

Yikes! What an insult to all the hard-working small-business owners out there. They didn't build their own businesses? The Romney campaign made some effective ads, with these remarks playing in the background, and small-business people describing how they struggled to get their firms going. The problem is, that quote above leaves some bits out—specifically, a few sentences before the last two. Here's the full transcript:

> If you've been successful, you didn't get there on your own. You didn't get there on your own. I'm always struck by people who think, well, it must be because I was just so smart. There are a lot of smart people out there. It must be because I worked harder than everybody else. Let me tell you something: there are a whole bunch of hardworking people out there.
>
> If you were successful, somebody along the line gave you some help. There was a great teacher somewhere in your life. Somebody helped to create this unbelievable American system that we have that allowed you to thrive. Somebody invested in roads and bridges. If you've got a business, you didn't build that. Somebody else made that happen.

Oh. He's not telling business owners that they didn't build their own businesses. The word 'that' in "you didn't build that" doesn't refer to the businesses; it refers to the roads and bridges—the "unbelievable American system" that makes it possible for businesses to thrive. He's making a case for infrastructure and education investment; he's not demonizing small-business owners.

The Obama campaign pulled a similar trick on Romney. They were trying to portray Romney as an out-of-touch billionaire, someone who doesn't know what it's like to struggle, and someone who made his fortune by buying up companies and firing their employees. During one speech, Romney said: "I like being able to fire people who provide services to me." Yikes! What a creep. This guy gets off on firing people? What, he just finds joy in making people suffer? Sounds like a moral monster. Until you see the whole speech:

> I want individuals to have their own insurance. That means the insurance company will have an incentive to keep you healthy. It also means if you don't like what they do, you can fire them. I like being able to fire people who provide services to me. You know, if someone doesn't give me the good service that I need, I want to say I'm going to go get someone else to provide that service to me.

He's making a case for a particular health insurance policy: self-ownership rather than employer-provided health insurance. The idea seems to be that under such a system, service will improve since people will be empowered to switch companies when they're dissatisfied—kind of like with cell phones, for example. When he says he likes being able to fire people, he's talking about being a savvy consumer. I guess he's not a moral monster after all.

False Cause

Post Hoc Ergo Propter Hoc

(Latin for After something, therefore because of that thing)

Just because something regularly follows another thing, doesn't mean that it is *caused* by that other thing. As the saying goes, correlation does not imply causation. The False Cause fallacy happens when someone mistakes correlation for causation.

For example, apparently Ice Cream sales and new cases of Polio (before the vaccine) were very closely correlated. Why? Because Ice Cream causes Polio? Luckily, no! I love Ice Cream and I'd be heartbroken to find out that it is the cause of such a tragic disease. It turns out (I'm told) that Polio cases showed up in warm weather more often and of course Ice Cream sales are very closely correlated with warm weather. Correlation does not equal causation!

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

Here's another fallacy for which people always use the Latin, usually shortening it to 'post hoc'. The whole phrase translates to 'After this, therefore because of this', which is a pretty good summation of the pattern of reasoning involved. Crudely and schematically, it looks like this:

[Math Processing Error]

This is not a good inductive argument. That one event occurred before another gives you some reason to believe it might be the cause—after all, X can't cause Y if it happened after Y did—but not nearly enough to conclude that it is the cause. A silly example: I, your humble author, was born on June 19th, 1974; this was just shortly before a momentous historical event, Richard Nixon's resignation of the Presidency on August 9th later that summer. My birth occurred before Nixon's resignation; but this is (obviously!) not a reason to think that it caused his resignation.

Though this kind of reasoning is obviously shoddy—a mere temporal relationship clearly does not imply a causal relationship—it is used surprisingly often. In 2012, New York Yankees shortstop Derek Jeter broke his ankle. It just so happened that this event occurred immediately after another event, as Donald Trump pointed out on Twitter: "Derek Jeter broke ankle one day after he sold his apartment in Trump World Tower." Trump followed up: "Derek Jeter had a great career until 3 days ago when he sold his apartment at Trump World Tower- I told him not to sell- karma?" No, Donald, not karma; just bad luck.

Nowhere is this fallacy more in evidence than in our evaluation of the performance of presidents of the United States. Everything that happens during or immediately after their administrations tends to be pinned on them. But presidents aren't all-powerful; they don't cause everything that happens during their presidencies. On July 9th, 2016, a short piece appeared in the Washington Post with the headline "Police are safer under Obama than they have been in decades". What does a president have to do with the safety of cops? Very little, especially compared to other factors like poverty, crime rates, policing practices, rates of gun ownership, etc., etc., etc. To be fair, the article was aiming to counter the equally fallacious claims that increased violence against police was somehow caused by Obama. Another example: in October 2015, US News & World Report published an article asking (and purporting to answer) the question, "Which Presidents Have Been Best for the Economy?" It had charts listing GDP growth during each administration since Eisenhower. But while presidents and their policies might have some effect on economic growth, their influence is certainly swamped by other factors. Similar claims on behalf of state governors are even more absurd. At the 2016 Republican National Convention, Governors Scott Walker and Mike Pence—of Wisconsin and Indiana, respectively—both pointed to record-high employment in their states as vindication of their conservative, Republican policies. But some other states were also experiencing record-high employment at the time: California, Minnesota, New Hampshire, New York, Washington. Yes, they were all controlled by Democrats. Maybe there's a separate cause for those strong jobs numbers in differently governed states? Possibly it has something to do with the improving economy and overall health of the job market in the whole country?

Hasty Generalization

A hasty generalization is just that: it's when one generalizes about a group of people or things or events, but one does so *too quickly* and without enough evidence or with too small of a sample.

Figure *[Math Processing Error]*: I know the last ten hawks have tried to eat your chicks, but don't be hasty to generalize to all of us hawks! (Image Credit: Otto Speckter in *Picture Fables*)

If I'm at the grocery store and grab two avocados that happen to be rotten, it would be rash of me to exclaim "What's up with this place!?!? All of the avocados in this store are rotten!" You need a randomized (so as to be hopefully representative) sample of avocados before you make a generalization about all of the avocados in the store.

Similarly with other generalizations you make. If you're jumping to conclusions about a whole group of things on the basis of interacting with or observing only a few of those things, there's a good chance that you're being hasty in your generalizing.

From: Knachel, Matthew, "Fundamental Methods of Logic" (2017).

Philosophy Faculty Books. 1. http://dc.uwm.edu/phil_facbooks/1

Many inductive arguments involve an inference from particular premises to a general conclusion; this is generalization. For example, if you make a bunch of observations every morning that the sun rises in the east, and conclude on that basis that, in general, the sun always rises in the east, this is a generalization. And it's a good one! With all those particular sunrise observations as premises, your conclusion that the sun always rises in the east has a lot of support; that's a strong inductive argument.

One commits the hasty generalization fallacy when one makes this kind of inference based on an insufficient number of particular premises, when one is too quick—hasty—in inferring the general conclusion.

People who deny that global warming is a genuine phenomenon often commit this fallacy. In February of 2015, the weather was unusually cold in Washington, DC. Senator James Inhofe of Oklahoma famously took to the Senate floor wielding a snowball. "In case we have forgotten, because we keep hearing that 2014 has been the warmest year on record, I ask the chair, 'You know what this is?' It's a snowball, from outside here. So it's very, very cold out. Very unseasonable." He then tossed the snowball at his colleague, Senator Bill Cassidy of Louisiana, who was presiding over the debate, saying, "Catch this."

Senator Inhofe commits the hasty generalization fallacy. He's trying to establish a general conclusion—that 2014 wasn't the warmest year on record, or that global warming isn't really happening (he's on the record that he considers it a "hoax"). But the evidence he presents is insufficient to support such a claim. His evidence is an unseasonable coldness in a single place on the planet, on a single day. We can't derive from that any conclusions about what's happening, temperature-wise, on the entire planet, over a long period of time. That the earth is warming is not a claim that everywhere, at every time, it will always be warmer than it was; the claim is that, on average, across the globe, temperatures are rising. This is compatible with a couple of cold snaps in the nation's capital.

Many people are susceptible to hasty generalizations in their everyday lives. When we rely on anecdotal evidence to make decisions, we commit the fallacy. Suppose you're thinking of buying a new car, and you're considering a Subaru. Your neighbor has a Subaru. So what do you do? You ask your neighbor how he likes his Subaru. He tells you it runs great, hasn't given him any trouble. You then, fallaciously, conclude that Subarus must be terrific cars. But one person's testimony isn't enough to justify that conclusion; you'd need to look at many, many more drivers' experiences to reach such a conclusion (this is why the magazine Consumer Reports is so useful).

A particularly pernicious instantiation of the Hasty Generalization fallacy is the development of negative stereotypes. People often make general claims about religious or racial groups, ethnicities and nationalities, based on very little experience with them. If you once got mugged by a Puerto Rican, that's not a good reason to think that, in general, Puerto Ricans are crooks. If a waiter at a restaurant in Paris was snooty, that's no reason to think that French people are stuck up. And yet we see this sort of faulty reasoning all the time.

9.E: Chapter Eight (Exercises)

Exercise $9.E.1$: Fallacies of Induction

Identify the fallacy of weak induction being illustrated by each. Remember that "no fallacy" is always an option. Try to decide whether this is good reasoning or not.

A. There haven't been any studies about the connection between this genetic mutation and breast cancer, so I choose to believe that it is that gene that caused my cancer.

B. If we don't stop them from outlawing bump stocks and large clips, it's only a matter of time before they'll take every single gun!

C. Nevermind that study showing a lack of correlation between this herb and increased health. I already know that it helps repair bones and prevent genetic damage, so that study must be faulty.

D. If we allow them to enforce drug laws in any way, pretty soon they'll decimate the community by jailing everyone for possession of even benign drugs.

E. I saw three kids outside the convenience store just up to no good. All kids these days have no respect.

F. I've been regularly riding my bike lately and I was just diagnosed with cirrhosis. I guess I have to stop riding my bike.

G. I know that aliens regularly visit us, so that strange experience I had last night must have been aliens.

H. I'm pretty sure that gravity explains why things tend to fall down toward the Earth, so that physical theory which suggests that it isn't gravity should be treated with suspicion.

I. We don't have great evidence about the beginning of the universe, so it must've started all at once with a big bang where a singularity exploded, creating time and space as well as all matter and energy.

J. Once we allow students to have some say in their education, they'll demand control over every aspect of their schooling and we'll lose control entirely.

K. I know he's a bad person, so the fact that he has been nothing but pleasant and kind and forthright to me for the past decade must be explained by the fact that he is trying to manipulate me.

L. I had an experience that I couldn't explain, so it must have been a ghost or a demon haunting me.

CHAPTER OVERVIEW

10: Ethical Reasoning and Evaluation

One of the most important things we reason about is *what to do*. How should we live our lives? What's a good person? What's the best life? What sorts of actions are out-of-bounds? What is required of us in situations of great need? What action is best in this situation?

These are *ethical* or *moral* questions.

We also reason about how to get things done, what the best means for attaining our goals are, what strategies work, etc.

These are *prudential* or *pragmatic* or *practical* questions.

We also reason about a lot of other things. Scientific questions, philosophical questions, legal questions, aesthetic questions, epistemic questions (questions about what we know and how we come to know things).

10.1: Ethics vs. Morality

There's no standard distinction between the 'ethical' and the 'moral.' Which are ethical questions? Which are moral questions? Who knows?

I like to think about them the following way:

The ethical (from Greek *ethos*) is a really broad category encompassing questions about everything we do. The ethical is about your relationship with yourself (and if you're a theist about your relationship with God).

The moral (from Latin *mores* or customs) is a narrower category encompassing only questions about our relations with one another. Moral questions are like the morality of abortion, murder, theft, lying, etc. They're about how we interact with other agents/actors.

> A sub-set of moral questions are *political*: how should we govern our society? What taxation schemes are fair/just/moral? What is a moral policing strategy? Etc.

On this conception, the ethical encompasses the moral and political because ethical questions are questions about the good life and what we ought to do, whereas moral questions are about what we ought to do to and with one another.

It's important to note, though, that this isn't an authoritative way to draw the distinction. There are other ways to do so. In this class, I tend to just use 'moral' and 'ethical' interchangeably.

10.2: Universalism vs. Relativism

Are moral values relative? If so, relative to what? Are they absolute/universal? If so, universal to what?

A lot of people are attracted to the idea that moral values are relative. But we don't quite know what that means until we dig a bit deeper (remember the second week, when we talked about meaning and definition? What do we mean by 'relative'?).

For instance, we might think moral values are relative to human beings or rational agents (porpoises, smart as they are, aren't subject to the same ethical constraints or rules). This seems really plausible. But it's the exact same thing as saying that moral values are *universal* to human beings or rational agents. Until we've filled in the blank in "universal to ____" and "relative to____", we haven't clearly said anything at all.

Figure *[Math Processing Error]*: Morality by H. Kopp Delaney

Some people think moral values are relative to culture: India has one set of values, the European west a different set of values; or maybe the American South (and most of rural California) has one moral standard and the American Coastal cities have a different moral standard.

Others think moral values are relative to individuals: I have my morality, and you have yours. Live and let live. There is nothing I can appeal to that we share, I can only appeal to my moral values.

On either version of relativism, we can't critique the moral values of others (perhaps other cultures, perhaps other individuals), there's no such thing as moral progress (every consistent set of moral values is totally incomparable), and we can't find any common moral values (which is clearly wrong since virtually all cultures and nearly all individuals abhor violent unnecessary killing, torture, rape, and other horrible things). With these considerations in mind, many conclude that relativism is wrong and that there is some set of moral values or principles that transcend individuals or cultures and extend to all human beings.

In this discussion, you may have noticed that I slid back and forth between two uses of the word 'morality': there's a use—the descriptive sense of the word—according to which a morality is any consistent set of principles that governs the actions of an individual or a community. So if we use morality in this way, it will make more sense to think that morality is relative: I have my morality and you have yours.

But if we think of morality in its *normative* sense, then it's harder to be a relativist about that kind of morality. According to the normative sense of the word 'morality', morality is a system of right and wrong that different individuals and different cultures debate about. What does morality require here? Is generally a question we use when we want to know what's right. Full stop. Not when we want to know what is right *for an American* or *for me*. Moral reasoning focuses on the normative sense of the word: what is right and what is wrong (not what principles govern my actions and which govern the actions of people I think are wrong).

10.3: Three Basic Systems of Morality

Any moral situation—any situation that morality governs—has three basic components:

1. **An Actor**: you, me, Jesus, Buddha, Malala Yusafzai, etc.

2. **An Act**: killing, helping, stealing, healing, saving, protecting, tormenting, etc.

3. **The consequences of the action**: People died, a person is alive that otherwise wouldn't have been, a person's burden was lighter, a family was protected, someone was caused a great deal of physical and psychological pain.

We have three basic systems which correspond to these three components. Whichever component of the moral action or situation you think is most morally relevant suggests which of these moral systems you tend to employ in moral reasoning.

- look at the consequences

2. **Deontologist**:

- look at the type of action: lying, helping, killing, healing, etc.

3. **Virtue Theorist:**

- look at the actor's character

Figure *[Math Processing Error]*: Bernardo di Stefano Rosselli 1450-1526

There are other ethical systems as I discuss in the lecture, but these are the main three. Buddhist ethics is a form of Virtue theory, but emphasizes different virtues than does Aristotle and has a different means of identifying what is virtuous and what is not (what is vicious or is a vice). Aztec ethics is also a virtue tradition, but it is one which emphasizes the community rather than individual character. Confucian ethics is similar: it's a virtue tradition, but emphasizes communal values. Karmic ethics is a "what goes around comes around" ethic, where at least one primary motivation for acting morally is that there are consequences to our actions. It might be in the next life or in the afterlife or the like, but bad actions have bad consequences that have a way of finding their way back to us some day. Same for good actions and good consequences. Divine command theory is the view that the ethical is whatever God commands us to do and the morally wrong actions are those we've been commanded not to do. Filling in the content here is tough, and must rely on some form of revelation, like from a Bible or Qur'an or the like. Finally, Care ethics is the view that what is important in ethical deliberations (decision making) is that we show proper care for those around us. Responding appropriately to the vulnerabilities of others is of paramount importance, and abstract considerations like Justice or Fairness or the like take a back seat. Arguably, care ethics is again a form of virtue theory. Care ethics arose out of a feminist response to male psychologists who thought it was obvious that someone was "less fully developed" who placed more emphasis on the immediate relational needs and less on abstract notions like justice and duty. They argued that these people are not less developed, but instead simply developed differently. This is an attempt to characterize the feminine approach to ethics—an approach which *tends to* come more naturally to women, but isn't necessarily associated with a particular gender or sex.

Let's go a little more in-depth on each of our three main theories...

Aristotelian Ethics or more broadly "Virtue Ethics"

Aristotle maintains that ethics is the study of how to be successful in life, how to flourish as a human being, and so what to do in life. Ethics is very broad. It's the study of how humans should act to be the best humans they can be.

Figure *[Math Processing Error]*: Aristotle is the dude on the right. The left? That's a dude named Plato. (Michelangelo's "The School of Athens")

With this in mind, Aristotle thinks that an ethical life is a life where doing what was right *comes naturally*. Just like an archer trains herself to shoot by repetition and the like so that when it comes time to shoot in the midst of a competition or battle it comes naturally to her, an ethical life is like training for a sport. We may need explicit guidance and rules early on in life, but after a while, the rules become less important and our own well-trained moral judgment becomes more important. Sport training similarly involves explicit guidance from a coach or trainer until we become so good that running with good form or aiming properly or throwing accurately just comes naturally—we no longer have to think about it.

What does it mean to train ourselves well? To develop *habits* to do the right thing and make the right judgments in the right circumstances. Habits, when we're talking about ethical or moral habits, are called *virtues*. We call it virtue ethics because the focus of the theory is on training ourselves well—on instilling virtues in ourselves and in others. We're trying to be people of good moral character, people to whom doing the right thing is second nature.

How do we know what the virtuous is? Easy, just find the mean between two extremes (this is Aristotle's somewhat controversial method for identifying the virtuous habits). If we have too little courage, we'll be cowards. If we have too much, we'll go running into battle without a plan and be killed immediately. Instead, we want enough courage to be able to do the right thing at the right time, and enough restraint to know when to hold back and bide our time—we want a mean (or middle amount) of courage. Similarly, we don't want to be lazy, but we also don't want to be ruthlessly ambitious—we want proper ambition. We don't want to be overly shy, but we also don't want to be shameless and incorrigible—we want to be modest. Hopefully you can get the basic idea here.

What's the easy test that an Aristotelian or virtue theorist can use to find out what to do in a given situation? How do we know what is right or promotes flourishing? One answer is that we can apply the "WWxD Test." What's that, you ask? It's taken from the phrase "What Would Jesus Do?" found on bracelets in the 90's, but instead of just Jesus, we put the variable x because the moral exemplar (the moral example we follow as a role model) might be someone different for different people. If you're Muslim, it's "what would Muhammed do?" If you're Buddhist, it's "what would Buddha do?" You simply replace the x with whomever you think is a moral exemplar—a virtuous person. Put that virtuous role model in your place and ask what they would do in the same situation. What sorts of things might they focus on? What might convince them to do one thing rather than another? What would they think are good reasons to act in one way or another?

Another virtue ethics test or heuristic is called the "Disclosure Test." Just imagine, before you do something, that everyone you know and love will find out that you did it. Would that be okay with you? For this test to work, as a disclaimer, you'll need to ignore or set aside things that are private and embarrassing and focus on things that are potentially public actions—actions that might be right or wrong. If you wouldn't want everyone to find out that you stole a bit of cash from work, then don't do it. If you wouldn't want everyone to find out that you've been harassing a coworker repeatedly, then don't do it. If your parents would be ashamed to find out that you got in a fight at school, then just walk away from the confrontation.

Buddhist virtue ethics and Jewish virtue ethics and... are all virtue traditions that simply put the emphasis on different virtues. Equanimity (being even-keeled and difficult to disturb from a state of inner peace) is a virtue for a Buddhist, but not for Aristotle, who focused on more social and political virtues. Study and obedience is a virtue for a Jew, but a Buddhist monk isn't so much concerned with study as with training. It's a bit more complicated than this in that there are more subtle differences between these groups of virtue theorists, but they're all close cousins.

Deontology/Kantianism

Deontological ethics focuses—instead of on virtues and moral character—on rights, duties, and general principles or rules. The focus here is on the act itself. Was it a murder? Then it was wrong. Same with lying, cheating, stealing, and torturing. No matter the intended consequences, no matter the character of the actor, no matter the actual consequences. Stealing is always wrong.

Diving right in, we can posit two special rules that will guide our actions as a Kantian deontologist (someone who follows the ethical theory of Kant). First, we should never treat our circumstances as exceptional. If there's a moral rule, then I must follow it no matter what. After all, I'm not special or worthy of special consideration just because I'm *me*. Instead, I must follow the rules like everyone else. Every time we act, in fact, we should pretend that we're broadcasting a new moral law to everyone in the world: in these circumstances, do this sort of thing! If we buy conflict-free diamonds, we're in some way affirming to the world that *everyone* should try to buy conflict-free diamonds. If we drive a gas-guzzling SUV simply because we like the look and feel of it and not out of any practical necessity, then we're in some sense telling the world that it's okay for everyone to do it. Remember: you aren't exceptional, the rules apply to you in the same way they apply to anyone else.

Second, you're supposed to treat everyone only as ends in themselves and never only as means to an end. In short: don't use people. Don't manipulate others into doing what you want them to do even though they wouldn't—if they stopped and thought about it—want to spend that time doing that. Don't lie to get other people to believe something which motivates them to help you with your projects when they themselves have their own projects to attend to. Again, you aren't special, so don't disrespect the values, goals, and projects of others in service of your own values, goals, and projects. For example, don't use manipulative sales techniques to try to get someone to spend more money than they need to on a product just so you can pocket the profits.

These two rules (along with a third we won't discuss) are called Kant's "Categorical Imperative." Nevermind now why it has such an odd name, but the idea is that we have one basic duty required of us as rational actors—as beings who act because of reasons they have for acting and not only because of urges and appetites. That duty is to treat other rational actors with the same respect with which you want them to treat you. From that basic duty come lots of individual duties—like don't lie, cheat, and steal. These "perfect duties" form the boundary lines of life: the "do not cross" lines. We can do as we like as long as we aren't lying, cheating, stealing, torturing, breaking promises, etc., etc. Duty doesn't require us to buy strawberry as opposed to cookie dough ice cream, but it does require that we *buy* rather than *steal* that ice cream.

Figure *[Math Processing Error]*: Mutual Respect, according to *American etiquette and rules of politeness*, an 1883 pamphlet

The easy way to think about it is the follow. The categorical imperative breaks down into two "tests" or "heuristics": every action must be universalizable and reversible. When I buy cookie dough ice cream, I'm saying it'd be okay if anyone in similar circumstances bought cookie dough ice cream. I'm not special, I'm doing something that anyone could morally do. So my action is universalizable. I can make a universal rule out of it: "go ahead, buy cookie dough ice cream." If I lie to my teacher to get out of taking a test, though, I can't be willing to universalize the rule. To say it's cool if anyone lies to their teachers to get out of taking tests is to say that education is a mere game we play and anyone can bend the rules however they like. It'd undercut the very reasons we go and get an education in the first place (aside from the fact that our parents want us to go). More evocatively, to cheat on a test is to undermine the very practice of taking a test. If everyone cheated, then the test would be useless and there would be no test to take in the first place since teachers would stop giving them. Similarly, if everyone lied all the time, then no one would ever believe anything that anyone said and therefore we couldn't really lie effectively since no one would believe us anyways.

Okay, so some actions are universalizable and others are not—we can't consistently extend the right to do the same sort of action to everyone else. Similarly, some actions are *reversible* and others aren't (just like some basketball jerseys are reversible and others are not, am I right?). If I hit you in the face for no reason, that is not a reversible action, since I wouldn't want you doing the same to me. If I help you get up after you've fallen off your skateboard in the park, then that action is clearly reversible because I would definitely want someone to help me up if I had fallen.

Kant went a step further than this. You not only have to do the actions demanded by duty (by the Categorical Imperative) but you also have to do them *because* they are demanded by the categorical imperative. You should act because acting in that way is your duty and not because it will really impress that cute guy or gal across the room or because it will bring you esteem or fame or financial backing. Easy for Kant to say, but it's a pretty difficult standard to live up to.

Consequentialism/Utilitarianism

Utilitarianism is a particular form of *Consequentialism*—the idea that the morally important aspect of an action is its *consequences*. For a Utilitarian consequentialist in particular, the idea is that the type of consequence we're interested in is *maximizing relative well being or happiness* (utility). We want each action to bring about the highest possible amount of well being relative to alternative actions we could have performed. I could travel across the world to deliver food to starving children, or I could simply wire money to the people already there who can get the food to them. Choosing between these two courses of action, I should weigh which one creates the most happiness or wellbeing.

It's a bit more complicated than that, though, since we also often *prevent* bad things from happening rather than only ever producing more or less happiness. We also produce unhappiness in the interest of producing more unhappiness, prevent good things from happening in the interest of preventing bad things from happening, etc., etc. It gets quite complex.

Either way, though, the idea is quite simple: an action is good if and only if it produces the most possible utility and/or prevents the most possible disutility relative to the alternative actions available to us at the time. If you want to know if an action is good, look at the alternatives and weigh the likely outcomes of each.

Figure *[Math Processing Error]*: A bison must die for the wolves to survive. The greatest good for the greatest number!

To complicate it even further, we should ask ourselves a question: should we act so as to *every single time* produce the most utility possible, or should we instead act according to a set of principles or rules that will—if we follow them consistently—produce the most utility? This is the difference between Act Utilitarianism—which says that you should do the calculation every time you act—and Rule Utilitarianism—which says that you should follow general rules which serve as signposts to guide you toward producing more utility than harm.

One more complication: are all pleasures or states of happiness equal? People on drugs are often quite happy, but the pleasure you get out of hiking in a beautiful park or grasping the unity and holistic beauty of a long solo in a jazz song surely are *higher* pleasures, right? Shouldn't we put more weight on the pleasure of attaining understanding of complex truths than of eating a chili dog? John Stuart Mill thought we should: higher pleasures, he argued, are "worth more" than lower pleasures and so our actions are better off for producing more higher pleasures than lower pleasures. Maybe, though, this is classist and Eurocentric nonsense. Maybe Beethoven isn't any better than Bachata even though it's more complex.

10.4: Levels of Evaluation

Your friend Amal is struggling. She's unmotivated to make positive changes in her life and is stuck in a job that pays too little and works her too hard. How do we explain what is going on with her?

Well, we have three basic options and the choice between them will determine how we *evaluate* Amal: is she a victim of circumstance? Is she *maladjusted* and in need of medical intervention? Or is she simply making bad choices?

The three levels of evaluation, as I like to call them are the different lenses through which we can view Amal's situation.

Material evaluation

At the **material** level of evaluation, we look at what sorts of *resources* Amal has access to and how Amal is put together as an organism. Is Amal's hypothalamus being impinged on by a benign tumor? If so, then the tumor is the best explanation for her lack of motivation—there's something *materially wrong* with her. Is she chronically malnourished? Does she lack the financial resources necessary to make the changes she'd otherwise make? If so, she's got a *material deficit*—not having enough resources of one kind or another is getting in the way of her doing what she'd like to do.

- "He's missing some capacity that the rest of us have"
- "If she had a little bit more resources, she could do great things."
- "He's got good genes, comes from a long line of distance runners."
- "It wasn't her saying all those things, that was the brain tumor."
- "They weren't ignoring you because they were mad at you, they were doing it because their brain was low on serotonin and so they were catatonic."
- "He was malnourished as a child, so now he has trouble with impulse control."
- "I'm sorry, I'm just really tired/in a bad mood/really hangry."
- "She didn't mean to do it, it was just a reflex!"

Agential evaluation

At the **agential** level of evaluation, we treat people like rational agents, who are responding to reasons for acting and making a more or less free choice. Agents are in charge of their actions, they have responsibility for the outcomes of those actions, and they are acting so as to express their values or reasons. Their own convictions and commitments determine what they end up doing rather than any material deficits or deficiencies. We spend a lot of time at this level of evaluation, especially when it comes to our negative evaluations of the actions and situations of other people and our positive evaluations of our own situations and actions.

- "She meant to do that"
- "They always go out of their way to make me feel welcome."
- "A lot of hard work and good choices got me where I am today."
- "You were ignoring me on purpose!"
- "What reason would someone have to do such a thing? What were they thinking?"
- "He knew what he was doing and he should take responsibility for his actions."

Structural evaluation

At the **structural** level of evaluation, we are interested in the contexts or structures in which people are embedded in evaluating their situations or choices. Why doesn't she eat healthier? Well, she lives in a food desert where the only food to buy for miles is at quickee marts and fast food restaurants. We're pointing to facts about the way society is ordered or the roles or situations people find themselves in to explain their behavior or situation. Why doesn't he study harder for his math exams? Well, his father left a few years ago, so he has to take care of his siblings after school and then work after dinner until midnight each night. It's a particular *role* that this character plays and all of the responsibilities that go along with it that explain why he does what he does.

Notice how we haven't talked at all about choices here. We're not interested necessarily in why people make the choices they do, because sometimes when we talk about the structures in which they're embedded, it becomes clear that they don't really have a choice. Other times structures merely make it so that someone's apparently irrational or imprudent choices are actually the best choice available. Here's an example: Sam works hard at their job, often picking up overtime and working extra shifts. Sam still, though, doesn't have any hope of saving for retirement or saving to buy a house, etc. They are stuck. So when Sam finally is able to save $200 in the bank after six months of skimping and saving, they face the choice of either a) keeping it in the bank so that in 30 years of scraping they can finally afford to get a more solid financial footing under their feet, or b) spend it on something nice for theirself now, so that they can do more than simply survive for the next 30 years. We could talk about personal responsibility here, but we could also recognize that the structures within society sometimes make a bad choice—buying a new phone or video game instead of saving for the future—the best available choice—living life while you have it so that you don't suffer your whole life waiting for a financially stable future that may never come. If a person's seemingly irresponsible choice is actually the best choice available to them, then we have good reason to think that a structural evaluation of their circumstances will be helpful in understanding their actions and life situation.

So when we engage in structural evaluation, we're evaluating a person's actions and circumstances insofar as they were caused by or constrained by the structures, contexts, systems, etc. in which that person is embedded and therefore also by reference to the roles that that person plays in those structures.

- "Of course they ended up doing that, they didn't have a choice!"
- "She never has time for brunch anymore because she's a mother."
- "Anyone in those circumstances would have done the same thing."
- "How can we hold someone responsible for doing what society has forced them to do?"
- "Growing up, we only had two options available to us: work at a fast food joint or deal drugs."
- "They live deep in rural California and so don't always have the easiest of access to fast internet."
- "She suffers from depression because she is constantly faced with hurtles that keep her from advancing at work and threats of violence and harassment when she's not at work."

When we're thinking at either the material or structural level, we're thinking about someone in a way that brackets or temporarily disregards their role in their life as responsible actors. Thich Nhat Hanh, a Vietnamese Buddhist Monk, gave the analogy of a head of lettuce growing in the ground. If it starts to shrivel or wither, we will think "it needs more water" rather than "it needs to make better decisions." If it is too small, we'll think "oh its genes code for less growth" or "maybe it didn't get enough sunlight." We won't think "that head of lettuce is at fault for being so small. If only it would've tried harder." Get it? By treating other people like heads of lettuce we ignore something important about them: that they have reasons for their actions and are often deliberately making the decisions they make. But we also open up a whole way of understanding their behavior that doesn't necessitate that they be at fault for everything bad that happens in their life.

A few upshots

First, we should note that we often think of ourselves as victims of circumstance or as people struggling to do our best with the resources at our disposal. We often, though, see other people as *agents*—as acting on purpose and for reasons. This is a pretty problematic way to go about our lives. We make excuses for ourselves while refusing to do so for others. Instead, I believe we should reverse this dynamic: we should see ourselves as agents and always hold ourselves accountable for our actions (except maybe give ourselves a break when we need it) and see other people as victims of circumstance or people determined to do what they do by habits, reflexes, and sometimes by the fact that they lack material resources.

A related point is that we're, empirically speaking, not good at this. We suffer from a cognitive bias called **fundamental attribution error**, where we tend to think of the actions of ourselves and of people in our "in groups" as excusable by appeal to circumstances or material resources/makeup; but we tend to think of the actions of people who aren't like us or are from an outgroup (even fans of a rival sports team, for instance) as *agents* who make bad choices and/or have bad intentions.

Our goal should be to *reverse* fundamental attribution error and instead be harder on those around us and strive to understand those who are not like us with as much empathy, understanding, and patience as we can muster. Try to see everyone else the way you see your best friend: if they screw up, you'll make an excuse for them or give them a break. Everyone messes up sometimes.

A different upshot is somewhat political: it's tempting to think that it is only our choices and hard work that determine our lot in life (where we end up in the social hierarchy), but if we take the material and structural perspectives seriously, it becomes harder to hold onto this absolutist view. We must admit, it seems like everyone would agree upon reflection, that sometimes circumstances or material resources determine how things go for someone. Some people are fantastically successful because their parents had a lot of connections (structural/relational perspective) and/or because they were given a large inheritance (material perspective) or maybe because their brains just naturally work well or their bodies are naturally fit and shapely or their hand-eye coordination is naturally fantastic without any training (still in the material perspective). The same goes for those who are unsuccessful. Life isn't *only* about the choices we make, but obviously our choices play a large role in shaping who we become and where we end up.

Finally, we have to be careful in choosing when to apply moral reasoning, or moralizing about a situation. A person may be simply making the best decision offered them by the circumstances or acting out an impulse brought on by the material structure of their brain. If that's the case, then moralizing or judging them on moral grounds might be inappropriate. We aren't always rational actors making rational decisions or acting for reasons; and furthermore what sorts of things a rational person would do changes dramatically with circumstances and context. Don't be too quick to judge other folks on moral grounds!

CHAPTER OVERVIEW

11: Mental Heuristics and Biases

11.1: Classical Economics

Our minds work in interesting ways. One of the main goals of this text is to get to know ourselves a bit better. One way to do that is to look at the weird way our minds are put together and the weird habits of mind that result. This chapter aims to explore what are sometimes called "cognitive biases"—or quirks of the way we naturally categorize and make sense of the world around us.

Economics is the study of the exchanging of goods. Classical economic theory sees humans who exchange goods (like money, cologne, gasoline, labor, time, etc., etc., etc.) as rational actors. The way to model human decision making, according to classical economic theory, is to treat each actor as faced with the problem of finding the most optimal action available.

Take Zito Here:

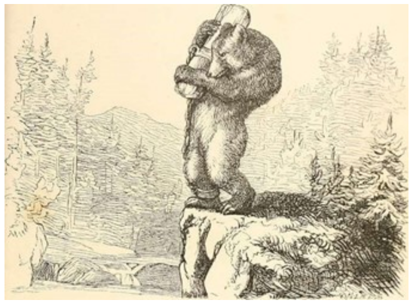


Figure *[Math Processing Error]*: Poor Zito has a log shackled to his leg. (Image Credit: Otto Speckter, *Picture Fables*)

Zito has a variety of choices available to him. He can drag the log behind him. He can carry the log around while walking on his hind legs. He can attempt to remove the shackle. He can throw the log in a river. And so on. Each of these options has an upside and a downside. A cost and a benefit. Here's what actually happens:

> BEAR AND LOG
>
> WELL, now, here is a wonderful thing!
>
> This great, huge log to my leg will cling.
>
> I'll get rid of you soon, I will;
>
> I'll carry you straight up yonder hill,
>
> And send you splashing, before I go, Into the river that runs below.
>
> The poor old bear he had reckoned wrong—
>
> The great log bore him with it along;
>
> Over and over he roll'd on the ground
>
> Till the brains in his head seem'd whirling round.
>
> He'd thought to free himself, but instead.
>
> He lay on the ground with the log, half dead.
>
> (Wilhelm Hey, *Picture Fables)*

Ouch. Zito appears to have made the wrong decision. He was faced with an optimization problem: the problem of finding the best available action, and he failed.

According to classical economic theory, the explanation is simple: Zito was being irrational. According to a more recent movement in psychology, social science, and what is now called behavioral economics, we shouldn't be so quick to judge. Especially if we see the same sort of thing happening over and over—people (or bears) making similar decisions that seem to be irrational; then we should investigate the possibility that they aren't be irrational. We should see if there's a general rule that is being applied to make these decisions and see if we can make sense of why those rules might in general be helpful—even if they aren't being helpful in the scenario that piqued our interest.

11.2: Heuristics

Zito appears to have followed a rule like the following: if something is bothering you, throw it off a hill into the river. This actually seems like a pretty good rule, right? If a badger bites onto your leg, throw it in the river. If a pinecone gets caught in your fur, throw it in the river. If a human sets up camp right next to the tree you like to use to scratch your back, throw them in the river. Yeah?

Of course in this particular case, this general rule doesn't seem to have worked.

A general rule like this—one that doesn't always work but which gets us where we need to go most of the time—is called a **heuristic**. That's a fancy word for a rule of thumb, a rough and ready strategy, or a shortcut. There's no promise that it's going to work *all* of the time. It just needs to work well enough to get you through the world in one piece. Rather than take all of the time and energy it takes to make the *best* decision, we can take a shortcut and use a heuristic to make a *pretty good* decision.

Daniel Kahneman—a behavioral economist—called heuristics "machines for jumping to conclusions." They allow us to quickly and easily make judgments about the world around us. Our minds, it turns out, work so darn efficiently because it is built around a series of heuristics. We use general rules to get by and that lets us make fast, efficient decisions when speed is what will keep us alive. It's a good way to build a mind because most of what a mind does is try to stay alive.

It's not so good a way to build a mind if what you're interested in is good reasoning, fairmindedness, or intellectual virtue. Heuristics all-too-often get in the way of thinking well.

The notion that has come to dominate behavioral economics is called "bounded rationality". Instead of being rational beings that always make optimal choices, we make the best choices we can given the resources we have to work with. We don't have the time, the energy, the knowledge, the motivation, or the processing power to be able to make perfectly optimal choices all the time. We instead make choices that are *good enough*. We make choices that *suffice* to *satisfy* our needs—this is called "satisficing" (I know, silly word, right? It's a combo of suffice and satisfy).

Let's look at a few Heuristics so we can get an idea of how they work:

Representativeness

The representativeness heuristic can be quite useful, but can also be the source of a lot of our most problematic thinking. The basic idea is that when faced with a new situation, we find the nearest prototype(s) in our mind and use what we know about that prototype to help us understand what is going on right in front of us.

If I see someone walk into a bank with a ski mask, then I look through my memories to see what most closely resembles the current situation before I settle on the prototypical bank robbery. I might even be able to make really good predictions based on this prototype. This person will lock the door behind them, knock out the security guard, shoot into the air, and then yell "everybody get on the floor!" This heuristic has been quite useful. It might even save lives.

Con artists exploit this heuristic all the time. They know that if they act like a particular prototype you have in your mind, you will associate certain things with them. If an older gentleman acts and dresses like your grandfather, then you might implicitly trust him (depending on your relationship with your grandfather, of course). You might even help him out of the financial bind he's in by loaning him some money…You get the idea?

This heuristic may also be behind a lot of our bigotry. We have racist, sexist, and ableist (among others) prototypes that tell us that when we see a person of a certain type, we should expect a certain thing. One can see how this might be problematic.

> **Here's the explanation of the Representativeness heuristic from Jason Southworth and Chris Swoyer's text "Critical Reasoning: A User's Manual". Shared under a CC BY-NC 4.0 license.**
>
> Mike is 6'2", weighs over 200 lbs., (most of it muscle), lettered in two sports in college, and is highly aggressive. Which is more likely?
>
> 1. Mike is a pro football player.
>
> 2. Mike works in a bank.
>
> Here, we are given several details about Mike; the profile includes his size, build, record as an athlete, and aggressiveness. We are then asked about the *relative frequency* of people with this profile that are pro football players, compared to those with the profile who are bankers.

What was your answer? There are almost certainly more bankers who fit the profile for the simple reason that there are so many more bankers than professional football players. We will return to this matter later in this chapter; the relevant point here is that Mike *seems* a lot more like our picture of a typical pro football player than like our typical picture of a banker. And this can lead us to conclude that he is more likely to be a pro football player.

Many of us made just this sort of error with Linda. Linda, you may recall, is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and she participated in antinuclear demonstrations.

Based on this description, you were asked whether it is more likely that Linda is (i) a bank teller or (ii) a bank teller who is active in the feminist movement. Although the former is more likely, many people commit the conjunction fallacy and conclude that the latter is more probable.

What could lead to this mistake? Various factors probably play some role, but a major part of the story seems to be this. The description of Linda fits our profile (or stereotype) of someone active in today's feminist movement. Linda strongly resembles (what we think of as) a typical or representative member of the movement. And because she resembles the *typical* or *representative* feminist, we think that she is very likely to be a feminist. Indeed, we may think this is so likely that we commit the conjunction fallacy.

We use the **representativeness heuristic** when we conclude that the more like a representative or typical member of a category something is, the more likely it is to be a member of that category. Put in slightly different words, the likelihood that x is an A depends on the degree to which x resembles your typical A. We reason like this: x seems a lot like your typical A; therefore, x probably is an A.

Sometimes this pattern of inference works, but it can also lead to very bad reasoning. For example, Linda resembles your typical feminist (or at least a *stereotype* of a typical feminist), so many of us conclude that she is likely to be a feminist. Mike resembles our picture of a pro football player, so many of us conclude that he probably is one. The cases differ because with Linda we go on to make a judgment about the probability of a conjunction, but with both Linda and Mike, we are misusing the representativeness heuristic.

Overreliance on the representativeness heuristic may be one of the reasons why we are tempted to commit the gambler's fallacy. You may believe that the outcomes of flips of a given coin are random; the outcomes of later flips aren't influenced by those of earlier flips. Then you are asked whether sequence HTHHTHTT is more likely than HHHHTTTT. The first sequence may seem much more like our conception of a typical random outcome (one without any clear pattern), and so, we conclude that it is more likely. Here the representative heuristic leads us to judge things that strike us as representative or normal to be more likely than things that seem unusual.

Specificity Revisited

We have seen that the more detailed and specific a description of something is, the less likely that thing is to occur. The probability of a quarter's landing heads is 1/2, the probability of its landing heads with Washington looking north is considerably less. But as a description becomes more specific, the thing described often becomes more concrete and easier to picture, and the added detail can make something seem more like our picture of a typical member of a given group.

In Linda's case, we add the claim that she is active in the feminist movement to the simple claim that she is a bank teller. The resulting profile resembles our conception of a typical feminist activist, and this can lead us to assume that she probably is a feminist activist. This may make it seem likely that she is a feminist activist. And this in turn makes it seem more likely that she is a bank teller *and* a feminist activist than that she is just a bank teller. But the very detail we add makes our claim, the conjunction, less probable than the simple claim that Linda is a bank teller.

In short, if someone fits our profile (which may be just a crude stereotype) of the average, typical, or representative kidnapper, scrap-booker, or computer nerd, we are likely to weigh this fact more heavily than we should in estimating the probability that they are a kidnapper, scrapbooker, or computer nerd. This is fallacious, because in many cases there will be many people who fit the relevant profile who are not members of the group.

Anchoring and Adjustment

Do you think more than 10% or fewer than 10% of North Koreans support a change in their leadership? Now, what do you think the actual percentage is?

Do you think more than 3 million or fewer than 3 million live in Wyoming? Now, what do you think the actual number is?

You might have guessed 20% of North Koreans, but you probably didn't guess 80% or even 60% or maybe even 40%.

You may have guessed that 1 million or 4 million live in Wyoming, but you probably didn't guess 300,000 or 30 million.

These "anchors" tend to keep us tethered around a particular range of answers even if we might have—without the anchor—guessed a much higher or much lower percentage. It turns out this is fairly robust in empirical studies: people tend to cluster their guesses around the anchors they are given.

This phenomenon is called "anchoring and adjustment" because we tend to anchor to the first piece of information we have about a new domain (even if it isn't presented as a fact) and then only "adjust" up or down from there. We don't tend to, when asked what the actual number is, wipe our minds clean of the anchor and start fresh. We tend, instead, to use the anchor as a clue to what the appropriate range is.

Consider this next time you go to pick out a new shirt or a paint color or a new significant other. Are you comparing them to the first shirt you saw, the first color that caught your eye, or your ex-partner? If so, you're probably adjusting from your anchor (your point of reference) rather than thinking in a fresh way about the decision.

Availability

> *"This is a mechanism that takes whatever information is available and makes the best possible story out of the information currently available, and tells you very little about information it doesn't have. So what you get are people jumping to conclusions. I call this a "machine for jumping to conclusions." -Daniel Kahneman*

Understanding that we operate according to the availability heuristic is one of the most important lessons we can learn, in my opinion. This underlies so much problematic reasoning that one could teach a whole class on the availability heuristic alone.

What I can think of, is all there is. What occurs to me in the moment is all I need to think of in order to make good judgments. What I can recall is much more likely than what I can't. At least that's what the availability heuristic would have you believe.

Let's say you're booking a flight (thanks Kendra Cherry for the example), and you remember a number of recent airline accidents, terrorist attacks, and disasters. Suddenly you are thinking that taking a train might be a better option.

Or maybe you're thinking about the representation of gay and lesbian couples on television, and you can think of a lot, so it seems like over half of the couples on TV these days are gay or lesbian (I once heard Glenn Beck make a similar claim on his radio show, of course not aware that he was in the grips of the availability heuristic).

Maybe you work in a hospital or police station or somewhere similar where the people you interact with are often up to no good—seeking pain pills to feed their addiction or being arrested quite often. Maybe the fact that you can remember so many of these people begins to flavor your idea of what it means to be poor or homeless or not neurotypical. You don't interact with the homeless folks who aren't being arrested, so you don't see them and consequently do not include them in your calculations of how many people who are homeless have been arrested before, or how many ER patients are seeking prescription drugs.

The availability heuristic runs on the quickness and vividness of recall. When I think of the US Congress, I think of Kamala Harris, Kirsten Gillibrand, Nancy Pelosi, Joe Biden, Mitch McConnell, Paul Ryan, and Claire McCaskill. Not as many men as women. In fact, when I do think of a man—particularly a white man—in congress, I might think of them as "another white male congressman" and not notice them as much because white men still make up the vast majority of members of congress.[1] So when I start to think of whether we have a representation problem in congress, I might not be so quick to think "yes" because I think of so many women right off the bat when I think of members of congress. The quickness of recall of the women members makes it seem to me like there are more women in congress than in fact there are.

The more *available* some set of examples are to your mind, that is, the more prevalent you think that phenomenon is. You see a car wreck and start to think it's quite common. You see Jaws and the vividness of your memories of the attacks makes shark attacks seem overwhelmingly likely.

You might see something on the news all the time—child abductions for instance—and then grossly overestimate the likelihood of your children getting abducted. After all, the countless children who never get abducted or even live in the same town as a child who has been abducted don't register to you—they don't seem important, so you don't log them in your memory banks.

As with the other heuristics, this can sometimes make our mental lives much easier, but it can also really get in the way of good reasoning about the world.

Availability and Online Algorithm Bubbles

One can imagine how problematic the Availability heuristic can be in a world where the information you have access to is determined by algorithms which are designed to ensure that you only see things that you want to see. Companies like Facebook, Google, Twitter, and the like all run on software that puts in front of you things that you are more likely to click—to like, to follow, to comment, to share, to engage with in some way. This software also responds to explicit instructions you send it telling it that you don't want to see posts from this person, or you don't want to see paid advertisements and other sponsored content from organizations and companies like this, or you don't like this particular post, etc. The result is what people call a "bubble," which is a curated and selective set of inputs that you see, sort of like your own mini reality—a different version of reality than the one your wacky aunt or uncle sees when they log on. So step one is to recognize that the internet often works this way: sites show you things that you want to see and tend not to show you things you don't want to see (where "want to see" is shorthand for "will likely engage with it and/or no report it to the algorithms for filtering out of your feed). You might not like everything you see, but you are being shown a curated and personalized version of online reality rather than an impersonal and universal online world.

What does this have to do with availability? Simple: If availability is the heuristic that says "what I see is all there is"—it's a process wherein the mind generalizes based on what is available rather than on what is likely objectively true—and if "what I see" is selective, then the availability heuristic will generalize based on selective data.

If I think a certain kind of person is a certain kind of way, and then I log on and see examples of that all over my feed or search results, and then I react accordingly thus bolstering the algorithms that gave me those examples in the first place; Then the availability heuristic is likely to make me generalize based on those examples rather than based on my more objective assessment of how prevalent that problem actually is.

It's almost like the availability heuristic and online bubbles were tailor-made for each other. They weren't, but they sure get up to a lot of trouble when they get together. The availability heuristic is a bit dangerous—use beware!

Upshots

We could spend all day learning about different mental Heuristics, but instead it will serve us to draw out some implications of what we've learned. If it's true (and it seems that it is) that we employ short cuts or heuristics to get around, make judgments, and choose between alternatives, then we're made in a way that's more *efficient* than it is *rational*. The availability heuristic actually gives us the right answer sometimes, but so much of the time it leads us astray, which might cause us to question whether it's a good cognitive strategy at all. It is, however, much more efficient than spending the time, energy, and mental resources to remember absolutely every instance of a phenomenon. If we want to know if sharks are dangerous, it is far more efficient to remember Jaws, and then not go swimming. This is pretty bad reasoning, though, since shark attacks are extremely uncommon.

One upshot is that we shouldn't "trust our guts" since we now know that our "guts" are often subject to some powerful influences that can lead our gut instincts to judge incorrectly more often than not. We should be skeptical of our own intuitions and convictions, because we never know exactly where they come from.

Another upshot is that these heuristics may result in dangerous thoughts that are worth thinking carefully about. If I only ever see Arabs, Pakistanis, and other ethnic groups playing terrorists on TV, then I may come to have a particular bias towards being wary of people from these ethnic groups. If I only ever see stereotypical African-Americans on TV playing into a narrative about what it means to be African-American, then I may come to believe that this narrative and that stereotype are accurate descriptions of reality.

After all, every ____ I can think of is _____, so it seems reasonable to me to think that all _____s are _____. This is the structure that Availability inferences seem to take. Think up some examples and then form a generalization or tell a story that makes sense of this really small set of examples you've been able to recall in the moment. The problem is that we're often only able to recall exceptional examples—exceptions to the true generalizations about the kind of thing we're thinking about.

For more information on Heuristics: OpenPsyc

For more information on biases, pitfalls, and traps, see: More Biases, Pitfalls, and Traps, by Southworth and Swoyer .

---

[1] https://www.washingtonpost.com/news/the-fix/wp/2015/01/05/the-new-congress-is-80-percent-white-80-percent-male-and-92-percent-christian/?noredirect=on&utm_term=.272d8d98fab8

CHAPTER OVERVIEW

12: Identifying Good Sources of Information

We live in a world positively saturated with information. Information technology—computers, networking and the internet, mobile networks, online platforms like social media and video and audio hosting platforms, and mobile devices like smartphones—has us embedded in a world where almost any piece of information is readily available and we are flooded on a daily basis with information in the form of social media posts, podcasts, videos, news articles, fake news articles, and countless other media.

One of the most troubling results of this is that it's not always easy to tell what is good information—to sort the reliable sources from the unreliable, the misled from the misleading from the well-informed, the thoughtful from the ideologue. Here's a short guide to some considerations to take into account as you attempt to engage with the deluge of information you are faced with each and every day.

12.1: Sources of Information

One should always start by considering the source of information. Do a little research. Are they merely a host for people to post their own essays like Medium.com? Is it someone's personal blog? Are they a satire website like the Onion, the Borowitz Report, or the Babylon Bee? Are they masquerading as a local news site? Are they a deeply ideological source that is pushing one particular political agenda like Breitbart, Occupy Democrats, Info Wars, the Jacobin, or US Uncut?

Look for independent verification that the source is a good source of information.

Finally, when you've found a good source of information, it's a good idea to stick with it, **but you must also continue to get information from a variety of sources**. Every source has its biases and blindspots and the best way to get a complete picture is to look at a variety of sources. Read the National Review (an overtly conservative publication) and the Atlantic (a skews-left publication), listen to NPR and the NY Times, watch PBS, ABC, NBC and local news, and in general avoid the most biased of talk radio, hyper-partisan news publications, and ideological podcasts. Also follow some international news sources as well! Consider Al Jazeera, BBC, and Reuters. A varied media diet is essential to avoiding getting duped!

Does it have a real author?

FactCheck.org's guide (linked below) has a number of good examples of articles that have fake authors or no authors at all. Check it out. Always good to check the author and understand their reputation a bit before investing too much credence in a news story.

Is it biased?

All news is biased in some way or another. News sources like CNN, NBC, and ABC are clearly biased towards the sensational. News sources like MSNBC and Fox News are clearly biased towards different political ideologies. Sources like NPR and PBS can be harder to identify their bias, but they do tend to have a bias towards a mainstream status quo ideology. They also tend to follow the "main news story" of the day, and that always has a bias towards the political, the sensational, and the economic.

The question is: can you easily identify the bias and then account for that in your assessment of the information they are given? Can you sort out the facts they are reporting from the assessments of those facts or are they all intermixed in an inextricable web?

Some sources I would say aren't news sources at all because you can't extricate the facts from the assessment of those facts according to a certain ideology. John Oliver's Last Week Tonight, Rachel Maddow, Fox and Friends, the Daily Show, Tucker Carlson, Sean Hannity, and so on are all such examples. These aren't news sources: they are sources of political commentary and media criticism. If it's a news source, it should be fairly clear when they are giving you information and when they are offering analysis or assessment of that information.

A lot of this comes down to a difficult distinction: that between facts and opinions or assessments or evaluations. It's not at all easy to know how to make this distinction and philosophers like me are even less sure about how to make this distinction than are other folks, but here's a first pass:

- A fact is a bit of information that forms the "common ground" or shared understanding of people of widely different ideologies and biases. It's an authoritative claim that most independent sources agree on.
- An opinion, evaluation, or assessment is a narrative or viewpoint on the facts that brings in ideology, values, and judgments.

I cringe a bit at even trying to make this distinction because it is, as I said, very difficult to make, but hopefully this gives us a good starting place. Fox News' Tucker Carlson will give you some facts, but will weave those facts into a narrative so that you're only getting certain facts and the facts you do get are tinged with ideology and evaluation. Rachel Martin on NPR's *All Things Considered*, on the other hand, will largely give you a set of facts about who did what when, who said what and in what context, and what experts are saying as an analysis of what was said and done. Occasionally, some evaluation and ideology sneaks in, but generally speaking, she is very good at "sticking to the facts".

One quick way to check for bias is to search for the headline on a search engine. Look at who is sharing this news story. Consider whose interests this story seems to serve. Search on social media like Twitter or Facebook and look at which "bubbles" this story is making the rounds in. This can generally be a good guide to which ideological direction this story might be slanted.

Is it thoughtful and honest?

One of my go to tests for whether a news source is reliable is simply this: do they report when they are wrong? NPR constantly reports its own mistakes. This makes me more apt to trust them as a reliable source of information. Lots of news outlets never admit they were wrong, or they scapegoat one particular person when a mistake is made and so they never have to take responsibility.

This may sound weird: sources making mistakes *increases* your trust in them? Yep! As long as they openly admit it and correct those mistakes. Then I know that in the future if they make a mistake, they will correct it as soon as they find out it was a mistake.

Another test is this: do they consider the possibility that their assessment is wrong? Do they consider the other side fairly? Do they look at competing narratives and weigh reasons to accept either narrative? Do they consider counterarguments to their analysis? If so, then they are less likely to be shoveling ideology down your throat. If not, then they are far more likely to be doing so.

Who funded it?

One way of identifying bias—particularly when it comes to science articles, studies, polls, and so on—is trying to find out who funded the study, poll, etc. If a study or poll was funded by a presidential campaign, then there's a good chance that it is almost useless—unless, that is, they use a third-party nonpartisan polling organization that is well-respected. If a study is industry-funded and has findings supportive of or friendly to that industry, then you might put your skeptic's hat on. We might be distrustful of a study that has a vested interest in finding a particular outcome.

Does it try to get you to distrust "the others"?

C. Thi Nguyen clarifies in an Aeon article (linked below) that an Echo Chamber is an especially problematic social structure in that it not only shows us a partial and incomplete picture of the world, but also causes us to mistrust sources outside the echo chamber. I've never heard anyone on NPR say "you won't hear this on any other news outlets" or "you can't trust other sources on this because we're the only ones with the inside scoop" or "everyone else has bought into the lie, but we're here to give you the straight truth." If you hear these sorts of phrases, there's a good chance that the narrative they are spinning is biased, incomplete, or simply made up.

12.2: Specific Stories or Information

Once you've decided you're dealing with a good source of information, there are a lot more questions to ask about the actual story, information, or narrative you're interested in. Here are a number of those questions.

Is it Current? Is it Local?

Check the timestamp: is this three years old? If so, is it still relevant? Be careful about a phenomenon called "Context Collapse", coined by Danah Boyd: everything on the internet and particularly on social media seems to be taking place in *my context* *right now*. If you see an article being shared that says "Our area" for instance, look at the original source and original poster. Is it actually *your* area? Or is it a different area that only looks like your area because it is taken out of its original context and shifted to your own local context. If you see an article being shared that says "unemployment on the rise" for another instance, check the timestamp. Is it current? If it's three years old, then unemployment might *not* be currently on the rise. Time is important to context as well.

Everything we share online seems to be relevant *here* and *now* even though it often isn't. We just have to do some digging to find out whether it is or not.

Similarly, if someone posted something about a cure being found to Coronavirus, but it was posted in 2018, then it's not COVID-19 that they're talking about! Or if the article is from March or 2020 and you're in August of 2020, then that cure was probably not all it was cracked up to be since it isn't widespread news many months later.

What are others saying about it?

Another way of safeguarding against being duped by fake news or otherwise false information is to look at the same news story or piece of information from multiple independent sources. Are multiple world governments confirming the same bit of information? Are multiple news outlets with independent sources reporting the same story? Has the story or claim been debunked by other sources? Can we trust the sources doing the debunking? Have you checked independent fact-checkers like Snopes.com, factcheck.org, Politifact.org or the like? Have you looked to see what reporters from a variety of sources are saying about the story or piece of information (Twitter is probably a decent place to find this out)? Have you looked at sources from a variety of ideological backgrounds to find their takes on the story?

Is it Plausible?

We can often independently assess whether something is plausible or not just using our "common sense" (I hate that phrase, but it's somewhat applicable here). Is it plausible that Hillary Clinton is involved in a child-trafficking ring that has headquarters in your local pizza place? Not really. Is it plausible that Bill Gates has microchipping technology so advanced that it can fit in a vaccine needle—something that outstrips any nanotechnology known to exist? Is it plausible that he actually wants to microchip everyone? To what end? Put your skeptic's hat on and consider whether these claims are plausible and then go forward with researching the claims further through independent reliable sources only when you've decided it's at least plausible enough to warrant further investigation.

Is it convenient?

If it fits too neatly with a particular ideological narrative about current events, politics in your society, or something similar; then it might just be too convenient to be true. Sometimes the truth really does fit a particular narrative, but the more neatly and tidily it does, the more skeptical we should be.

Is it possible that it's a Deepfake?

Some information is just fake: it has been created from whole cloth to try to support a particular narrative or ideology. It's easy to make up quotes, but there is now the technology to create video and audio using machine-learning technology that is surprisingly convincing. It just needs some source data—like a huge amount of videos of Obama's speeches—and then some input data—like a person acting like Obama saying something outrageous—and then it can create a new video that looks and sounds convincingly enough like Obama saying the outrageous stuff. They can create images, voices, and even videos and at some point it very well may become impossible to tell which videos are fake just by looking at them.

What then? Well, we may be able to rely on alternative verification for videos. Maybe reporters will give affidavits stating that they were there when the video was shot. Maybe there will be an unfakeable piece of data embedded in real videos. Maybe the shear amount of videos people are likely to take at important events (whether it be news outlets or people on cell phones) will create a sort of public record. These are all questions for another day.

For now, though, remember that the more outrageous or convenient a video is, the more skeptical we should be that it is a genuine video.

12.3: Sources for this chapter

Reading and Guides

The Verge's guide to Fake News

Berkeley resources and links about Fake News

Fact Check guide to Fake News

C. Thi Nguyen on Echo Chambers

Snopes' list of Fake News Sites

Reveal's Guide to Fake News

Center for News Literacy

CBC Guide to Fake News

- CBC Guide to Misinformation and Disinformation

Southworth and Swoyer:

- Evaluating Sources of Information
- The Internet – Finding and Evaluating Information Online

Relevant Podcast Episodes

Radiolab on Deepfakes

NY Times' Rabbit Hole

Hidden Brain on Fake News: an Origin Story

NPR Lifekit on Fake News

On the Media: How Fake News Gets Made

A Youtube course on Media Literacy

CHAPTER OVERVIEW

13: Conclusion - Intellectual Honesty, Intellectual Humility, and Charity

> **Note**
>
> A lot from this conclusion has been copied into the chapter on Intellectual Virtues and Vices. I decided to keep it here in the conclusion as a means of emphasizing the centrality of these ideas to thinking well: thinking well is thinking virtuously. So here it is, some of it repeated, for emphasis:

This textbook and the courses associated with it are aimed at a number of different things. I and your instructors want to teach some basic skills in recognizing arguments, different types of arguments or inferences, different problems in reasoning, and so on. We want you to walk away from the course with some knowledge about arguments, argumentation and reason and with some skills in putting that knowledge to use.

More than that, though, I want to instill in my students a set of habits and dispositions—particularly intellectual habits or habits of mind. The three that come to my mind first are intellectual honesty, intellectual humility, and charity.

Intellectual honesty is the disposition to be truthful and sober in your assessment of your own knowledge. It's easy to claim that we know things and even to have confidence in what we know, but often we find that on reflection we *shouldn't* have as much confidence as we do. Confidence is cheap. What is of higher worth is the ability and disposition to recognize the things we don't know or shouldn't be confident in and the things that we do know and do have reason to be confident in. Much of what we think we know we think we know really because we read a headline while scrolling through Facebook or Twitter or someone told us once sort of off-handedly. These, when we think about it, aren't very good sources of knowledge. They aren't really grounds or justifications for our beliefs—or at any rate aren't very good justifications for our beliefs. Intellectual honesty is the disposition to take a beat, think about why it is that we feel confident in a belief and feel ready to assert it, and then proceed with a more honest assessment of what we know and why we think we know it.

Intellectual humility goes hand in hand with intellectual honesty. What is means to be intellectually humble, though, is slightly different from being honest. Intellectual humility is a disposition to recognize that even when we have good grounds for knowing something, there might always be something that upsets that understanding or set of beliefs. To be intellectually humble is to remember that human beings have been very confident many times in the past and often for very good reason, but have turned out to be wrong due to some false assumption somewhere in their thinking. It's the disposition to say "even if I have really good reason to believe what I believe, I still might be wrong."

Both of these are dispositions worth cultivating. But there is one more worth coming back to: what is Charity? To be charitable is to attribute the best intentions and strongest justifications to someone else. To interpret a set of actions charitably is to try to see those actions in terms of the most reasonable set of motivations or intentions behind them. To interpret someone's beliefs charitably is to attribute moral innocence to them as a person as far as is possible so as to give them the strongest possible benefit of the doubt. Only when you have really good reasons for doing so might you think of someone else as irrational, vicious (in the sense meaning the opposite of virtuous), or petty. Charity, then, is a habit of interpreting actions and beliefs in a good light— a rational and moral light.

All of these dispositions have their appropriate limits, of course: many beliefs and actions are just wrongheaded or irrational or bigoted and we needn't bend ourselves in pretzel knots trying to interpret them charitably. Many of our own beliefs are things we have really good reason for believing, so we don't need to be so humble that we refuse to believe anything. Some of us, furthermore, are really in a better position to know things and to reason about them. A false sense of humility stops being honest at a certain point.

Nevertheless, the hope is that you leave the class associated with this textbook a bit more skeptical, a bit more incredulous (more reticent to accept new statements as statements of fact), a bit more self-reflective. Circumspection is the act of looking back at oneself and at one's own beliefs and assertions. The hope is that you leave the course with a good habit to circumspect: to examine your own beliefs, assertions, and actions from a more disinterested stance. The tools in this text help move you one step closer towards this goal.

Index

D

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Thinking Well - A Logic And Critical Thinking Textbook 4e (Lavin)

Webpages: 81

All licenses found:

- CC BY 4.0: 97.5% (79 pages)
- Undeclared: 2.5% (2 pages)

By Page