Foundations

RJC

January 22, 2026

Think about something you think you know[1] about I.R. This could be anything—say,

1. a particular event, such as "Russia invaded Ukraine in 2014;" or
2. a particular description, such as "The United States is a democracy;" or
3. a stylized fact,[2] such as "civil wars occur less often in developed states;" or
4. a definition, such as "a state is a democracy if and only if its leaders are chosen by free and fair elections;" or
5. whatever.

Now, here's a question: what is it that you can do with this knowledge?[3] Several things come to mind,[4] though surely this list is not exhaustive:

1. You can communicate it to others. The examples above come in the form of language, in this case written plain-language English. That's a device for communicating knowledge.
2. You can use it to make predictions about the future. For example, perhaps knowledge that Russia invaded Ukraine in 2014 would help you predict that Russia would invade Ukraine again in 2022.
3. You can use it to theorize. For example, perhaps knowledge that civil wars occur less often in developed states would help you theorize about the causes of civil war. You might theorize that civil wars are caused by low opportunity costs[5] of rebellion, and that developed states have higher opportunity costs of rebellion.
4. You can use it to make decisions. For example, perhaps knowledge that the United States is a democracy would help you decide whether to move to the United States.

Notice how these uses of knowledge are related to one another. For example, you might use knowledge to make a prediction, and then use that prediction to make a decision you later justify to someone else. And your predictions may be rooted in theory. Regardless, the point is that knowledge is useful.[6] That probably ought to be clear.

[1] There are deep epistemological questions about what it means to know something. Let's leave those aside for now.

[2] A *stylized fact* is a generalization about the world that is widely accepted as true, but that is not necessarily true in all circumstances. The example given here is indeed a stylized fact; developed states experience civil wars less often than developing states, but they do experience civil wars.

[3] This is a question about the *instrumental value* of knowledge. The *instrumental value* of something is the value it has in helping you achieve some other goal. For example, the instrumental value of a hammer is the value it has in helping you build a house. The instrumental value of knowledge is the value it has in helping you achieve some other goal. But please do remember that knowledge has *intrinsic value* as well—it is valuable in and of itself.

[4] Of course, the most important thing you can do with knowledge is *sound cool at dinner parties,* but the other items on the list would seem paltry in comparison.

[5] An *opportunity cost* is the value of the best alternative that is foregone when a choice is made. For example, if you choose to go to college, the opportunity cost of that choice is the value of the best alternative to going to college—say, the income you could have earned if you had chosen to work instead. In the case of civil wars, perhaps people without access to high-paying jobs are more likely to join rebel groups because they have lower opportunity costs of rebellion.

[6] For *this,* you're paying.

Your humble professor—henceforth Y.H.P.—is of the opinion that knowledge is a little more useful when it's stored reliably.[7] Our communication becomes more precise, our predictions more accurate, our theories more robust, our decisions more informed. And it turns out that there exist tools that help us store things reliably. This chapter introduces some of those tools so that we can use them throughout the rest of the course.

[7] "Did he just go third person?" Oh yeah he did. "Is he gonna keep doing that?" Oh yeah he is. "Why?" For effect, and because it's fun. "How about these Gaffigan-esque call-and-response things?" Yep. "Are they gonna get old?" Aren't they already?

Here we will cover four concepts at an introductory level:

jeez louize this guy and the lists SHUT UP

1. *logic,* which allows us to communicate reliably;
2. *sets,* which allow us to gather things reliably;
3. *relations,* which allow us to describe things reliably; and
4. *functions,* which allow us to link things reliably.

You'll notice that these sound like things one studies in a math class. That's because they are—but don't freak out! None of this is all that mathy, and it's all remarkably useful for talking about I.R. You'll soon enough see that these are just tools, not anything to be afraid of. The experience ought to be like learning a new language, or how to play an instrument perhaps, or even how to use a new software program. It's a little weird at first, but you'll get the hang of it. But now I'm dawdling with pep talks when I should be getting on with the show. *We begin!*

Logic

Have you heard of the democratic peace thesis?[8] You can write a crude version of it like this:

> Democracies do not fight wars with one another.

[8] You sometimes hear it called the "democratic peace theory," but that doesn't sound right to Y.H.P.. "Theory" is a word that gets thrown around a lot, and it's not always clear what it means. Like, Y.H.P. certainly wouldn't say "Kant's theory of democratic peace" the same kind of thing as "Einstein's theory of gravity." But he could be convinced otherwise. Regardless, the one thing everybody can agree on is that democracies have not fought wars with one another. Go crush a dinner party with that one.

Now, that sentence is a *claim* about the world, and it could be `TRUE` or `FALSE`. Suppose you had a hunch that it was `TRUE`, but that you didn't quite know why. You might want to *justify* your belief to yourself.[9] You might say something like this:

> Democracies do not fight wars with one another because they are more likely to resolve their disputes peacefully.

[9] This is a good thing to do. It's called *epistemic justification,* which helps us understand why we ought to believe something based on the evidence we have. It's a philosophical concept, but one needn't be a philosopher to profit from its application.

Or this:

> Democracies do not fight wars with one another because their threats are more credible.

Or this:

> Democracies do not fight wars with one another because they are more likely to be allies.

That is, you might try to *explain* why the democratic peace thesis is `TRUE` by way of a *causal mechanism.*[10]

[10] A *causal mechanism* is a process that explains why a cause leads to an effect. For example, the causal mechanism that explains why democracies do not fight wars with one another is that they are more likely to resolve their disputes peacefully.

In the opposite direction, you might try to *assume* that the democratic peace thesis is TRUE and then use it to make a *prediction* about the future. For example, you might say something like this:

> Because they are both democracies, the United States and Canada will not fight a war with one another.[11]

It's not hard to imagine using reasoning like this to make decisions—analysts do that all the time, right?

To reason reliably—in either direction—we need a *language* that allows us to communicate, justify, explain, predict, and decide. That language is *logic*, and it's the first tool we'll introduce.

[11] This is a prediction that has been borne out by history. The United States and Canada have not fought a war with one another since the War of 1812. Nevertheless, Y.H.P. keeps an eye on his neighbors to the north—they might just be playing the long game. Except for Sidney Crosby; that guy's a saint.

STATEMENTS ARE THE BUILDING BLOCKS OF LOGIC. Consider the simple

> The United States is a democracy.

Observe:

1. This is the kind of sentence that can be TRUE or FALSE; it is what you might call *truth-apt*.[12] Such sentences are sometimes called *falsifiable*, a concept of supreme relevance in 20th Century philosophy of science.[13]
2. It would be impossible for this sentence to be both TRUE and FALSE at the same time; it is what you might call *bivalent*.
3. Were we to agree on a definition of "democracy," we could determine whether this sentence is TRUE or FALSE; it is what you might call *unambiguous*.[14]

And this is how we define statements.

---

1 Definition (Statements)

A statement *is a sentence that is unambiguously* TRUE *or* FALSE*.*

It bears repeating that FALSE statements are statements precisely because they could be FALSE: "The United States is not a democracy"—which I think most would agree is FALSE—is a statement just as much as "The United States is a democracy" is a statement. These are well-posed in ways that pathological utterances[15] like

> *What is the meaning of life?*
> *Go to the store.*
> *'Twas brillig, and the slithy toves did gyre and gimble in the wabe*
> *This class is the best class ever.*

are not. The first is a question, the second a command, the third two lines of nonsense poetry,[16] and the fourth a matter of opinion. These are not statements, as they are not unambiguously TRUE or FALSE. Fun as they are, they are not the stuff of logic, and they certainly aren't reliable enough for us to build up a theory of I.R. with them.[17] So, we'll just focus on statements for the rest of the class.

[12] Per Oxford Reference, "A sentence is truth apt if there is some context in which it could be uttered (with its present meaning) and express a true or false proposition. Sentences that are not apt for truth include questions and commands, and, more controversially, paradoxical sentences of the form of the Liar ('this sentence is false'); or sentences ('you will not smoke') whose apparent function is to make an assertion, but which may instead be regarded as expressing prescriptions or attitudes, rather than being in the business of aiming at truth or falsehood."

[13] It also so happens that many, if not most, political scientists try to focus on falsifiable claims. This is because falsifiable claims are the ones that can be tested. There is a lot of variety in PhD programs in political science, but nearly *everybody* has a working sense of falsifiability.

[14] Yeah, good luck landing on a consensus definition of "democracy." However, for the purposes of an academic discussion like this one, we can agree on a definition and proceed—or at least we can pretend like we could agree on a definition and proceed.

[15] BAND NAME I CALLED IT

[16] Uncle Lew, doing his nonsense poetry thing! Uncle LEW! Wubba lubba dub-dub!

[17] Apparently we're going to be building up a theory of I.R. in this class. Who knew?

Let's start with the simplest statements. Consider again:

The United States is a democracy.

You couldn't break that down into smaller statements. The fancy term for this is *atomic,* as in "statements that can't be broken down further will be the atoms of our logic." Notice that a statement could be about multiple things whilst remaining atomic: consider

The United States and Canada are not at war with one another.

This is a statement about two things, but it's still atomic.[18]

Now consider the really obnoxious exercise of gathering together a few hundred statements, such as

Afghanistan is a democracy.
Albania is a democracy.
—and so on, up until—
Zambia is a democracy.
Zimbabwe is a democracy.

Notice that they all have the same structure, namely

— is a democracy.

That is, these statements vary in their subjects, but they all have the same predicate.[19] It would be nice if we could just have a placeholder for the subject—a *pronoun,* if you will—and then just fill in the blank with the subject of our choice. So, we might say something like

s is a democracy.

where s is a placeholder for the subject of our choice. We call s a *variable,* which just means it's a symbol we can use as a stand-in for something else.[20] Thus, in case s = Afghanistan, then our statement becomes

Afghanistan is a democracy.

Or, in case s = Zimbabwe—which it's allowed to do, because that's what variables do, is vary!—then our statement becomes

Zimbabwe is a democracy.

We therefore allow sentences like "s is a democracy" to be statements, too.[21] We might use multiple variables in a single atomic statement, like

s_1 and s_2 are allies.

where s_1 and s_2 are placeholders for the subjects of our choice. In case s_1 = United States and s_2 = Canada, then our statement becomes

The United States and Canada are allies.

Notice that nothing precludes $s_1 = s_2$ = United States, yielding

The United States and the United States are allies.[22]

Finally, nothing precludes multiple occurrences of the same variable, like

s_1 and s_2 fought a war that s_1 won.

So despite their simplicity, atomic statements can pack a punch!

[18] Statements are distinguished by the number of things they are about, so thate statements about zero things, one thing, two things, and so on are all different kinds of statements. That means that our plain-language usage of terms might be misleading. Consider, for example

The United States and Mexico are allies.
The United States, Mexico, and Canada are allies.

Both statements end in "are allies," but they are different kinds of statements by virtue of the number of things they are about. This suggests that bilateral alliances are different in kind from trilateral alliances are different in kind from. . . .

[19] In fact, the term for such a logical device is indeed *predicate.*

[20] This is not how the term "variable" is used in other parts of political science. For example, we're not talking about the kind of variables one uses in statistics, nor are we talking about the kind of variables we'll use later on to *parameterize* models of I.R.

[21] Implicit in this exercise is some underlying universe of things that s can be. It would make little sense to say something like "Fred Rogers is a democracy," because Fred Rogers is not a state. The way that Fred Rogers isn't a democracy is different from the way that North Korea isn't a democracy, wouldn't you say? So in this example, the universe of discourse is states.

[22] Do you think this suffers from the same kind of `DOES NOT COMPUTE` problem that "Fred Rogers is a democracy" does?

It's kind of boring to walk around saying things like "s_1 is a democracy" and "s_1 and s_2 are allies."[23] It sure would be nice if we could take these atomic statements and turn them into, you know, like, *molecular* statements.[24] We can do that with *connectives,* which are logical devices that allow us to turn old statements into new ones. I'm going to show you seven of them. In the table given here, φ and ψ are placeholders for statements, and x is a placeholder for a variable.[25] Thus, in writing something like $\varphi \wedge \psi$, we mean that if φ = The United States is a democracy and ψ = Canada is a democracy, then $\varphi \wedge \psi$ = The United States and Canada are democracies. The same goes for the other connectives.

[23] my name is rob i am forty three years old i like potatoes

[24] I don't think this is a term.

Name	Symbol	# Inputs	Meaning	Example		
Negation	$\neg$	1	NOT	$\neg\varphi$	$\leftrightarrow$	"NOT φ"
Conjunction	$\wedge$	2	AND	$\varphi \wedge \psi$	$\leftrightarrow$	"φ AND ψ"
Disjunction	$\vee$	2	OR	$\varphi \vee \psi$	$\leftrightarrow$	"φ OR ψ"
Conditional	$\Rightarrow$	2	IF-THEN	$\varphi \Rightarrow \psi$	$\leftrightarrow$	"IF φ THEN ψ"
Biconditional	$\Leftrightarrow$	2	IF-AND-ONLY-IF	$\varphi \Leftrightarrow \psi$	$\leftrightarrow$	"φ IF-AND-ONLY-IF ψ"
Universal quantification	$\forall$	1	FOR ALL	$\forall x \varphi$	$\leftrightarrow$	"FOR ALL x, φ"
Existential quantification	$\exists$	1	THERE EXISTS	$\exists x \varphi$	$\leftrightarrow$	"THERE EXISTS x SUCH THAT φ"

[25] *Do not freak out.* Greek letters don't have to be scary. They are just symbols that stand in for statements or variables. Or fraternity mottos.

LET'S START WITH NEGATION. The negation of a statement is just the "opposite" of that statement. For example, the negation of "The United States is a democracy" is "The United States is not a democracy." Now, what does that mean,[26] the fact that one statement is the negation of another? Think of it this way:

[26] In logic, the question "what does that mean?" is a *semantic* question. It's a question about the meaning of a statement, usually with respect to some notion of truth. That's what *I* mean by "what does that mean?"

1. "The United States is a democracy" is a statement, and thus it is either TRUE or FALSE, but not both.

2. Similarly, "The United States is not a democracy" is a statement, and thus it is either TRUE or FALSE, but not both.

The fact that these statements contradict one tells us something about the patterns of truth between them: if one is TRUE, the other must be FALSE, and vice versa. We can see these patterns of truth among statements in a *truth table,* which is a table that shows the truth value of a statement for every possible combination of truth values of its constituent parts.

φ	$\neg\varphi$
FALSE	TRUE
TRUE	FALSE

Table 1: Truth table for negation.

The truth table for negation is given in Table 1. It has two rows, since the input φ can either be TRUE or FALSE. We call the rows of a truth table *cases,* so that we might say things like "In case φ is TRUE, $\neg\varphi$ is FALSE." For example,

> In case Mali is a democracy, Mali is not not a democracy.[27]

You're not going to win a Nobel Prize for that one, but it's still true.

[27] This demonstrates an important property of double negation, namely that $\neg\neg\varphi = \varphi$. Can you show this using a truth table?

Now let's move on to conjunction. Consider the statement

> The United States and Canada are democracies.

Under what conditions is this statement TRUE? Put differently, what about this molecule's atoms must be TRUE in order for the molecule to be TRUE? Well, both of its atoms must be TRUE, right? That is, both "The United States is a democracy" and "Canada is a democracy" must be TRUE in order for "The United States and Canada are democracies" to be TRUE.

We can see this in the truth table for conjunction, given in Table 2. It has four rows, since the inputs φ and ψ can each either be TRUE or FALSE in four different ways. We can see that "The United States and Canada are democracies" is TRUE only in the case that "The United States is a democracy" is TRUE and "Canada is a democracy" is TRUE—that's the last row of the table. Otherwise, it's FALSE.

φ	ψ	$\varphi \wedge \psi$
FALSE	FALSE	FALSE
FALSE	TRUE	FALSE
TRUE	FALSE	FALSE
TRUE	TRUE	TRUE

Table 2: Truth table for conjunction.

Next up is disjunction. Consider the statement

> The United States or Canada is a democracy.

Again, what about the atoms of this molecule must be TRUE in order for the molecule to be TRUE? This time, it seems like only one of the atoms must be TRUE, right? That is, either "The United States is a democracy" or "Canada is a democracy" must be TRUE in order for "The United States or Canada is a democracy" to be TRUE.

We can see this in the truth table for disjunction, given in Table 3. Since it again links two statements, it again has four rows. We can see that "The United States or Canada is a democracy" is TRUE in the case that "The United States is a democracy" is TRUE or "Canada is a democracy" is TRUE—that's the second, third, and fourth rows of the table. So, the only time "The United States or Canada is a democracy" is FALSE is when "The United States is a democracy" is FALSE and "Canada is a democracy" is FALSE—that's the first row of the table.[28]

φ	ψ	$\varphi \vee \psi$
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	TRUE

Table 3: Truth table for disjunction.

[28] This is an important point: "or" is *inclusive*. That is, "The United States or Canada is a democracy" is TRUE if *either* the United States *or* Canada is a democracy, *or both*. We often mean "or" in an exclusive sense—either you'll study for the exam or you'll fail the course—but that's not how it works in logic.

Notice that we can iterate "and" and "or" to link more than two statements. For example, suppose we wanted to say that the United States, Canada, and Mexico are all democracies. For any country s,[29] write $\mathbf{D}(s)$ to mean "s is a democracy." Then we could write

$$\mathbf{D}(\text{United States}) \wedge \mathbf{D}(\text{Canada}) \wedge \mathbf{D}(\text{Mexico}),$$

which encodes the statement that the United States, Canada, and Mexico are all democracies. If we had many statements—like $\mathbf{D}(s_1), \ldots, \mathbf{D}(s_n)$, where n is some counting number—we could write

$$\bigwedge_{i=1}^{n} \mathbf{D}(s_i) = \mathbf{D}(s_1) \wedge \cdots \wedge \mathbf{D}(s_n),$$

which encodes the new mega-statement that all of the underlying $\mathbf{D}(s_i)$s are TRUE.[30] How would these statements look in a truth table?[31]

[29] Here we are again using the letter s as a variable; it stands in for any state, so that it could be s = United States or s = Canada or s = Mexico or s = whatever. Remember, what variables do is vary!

[30] Similarly, we could write

$$\bigvee_{i=1}^{n} \mathbf{D}(s_i) = \mathbf{D}(s_1) \vee \cdots \vee \mathbf{D}(s_n),$$

which encodes the new mega-statement that at least one of the $\mathbf{D}(s_i)$ statements is TRUE.

[31] A good first step is the following: if I have n underlying statements, how many rows are in the corresponding truth table? What does that tell you about the number of cases?

NEXT UP IS CONDITIONAL. This is really important, because it approaches—but does not fully characterize!—*causal* relationships and *hierarchical conditions.* Consider

> If the United States is a democracy, then it holds elections.

This seems to convey something a little deeper than just the "opposite" statement would, or and- or or-linked statements would. It seems to convey that your knowledge about some quality of the United States—namely, that it is a democracy—is *sufficient* to make a prediction about some other quality of the United States—namely, that it holds elections.[32]

Now, what would make this statement FALSE? That is, what would have to be the case about the United States in order for this statement to be FALSE? Certainly, the statement is FALSE in case the United States is a democracy and it does not hold elections; that means that the reliable prediction promised by the conditional is not borne out. Of course, when both statements are TRUE, the conditional is TRUE, too—that means the reliable prediction is borne out. But what about the case where the United States is not a democracy? If Y.H.P. said to you "If the United States is a democracy, then it holds elections," and you knew that the United States was not a democracy, would you say that he was FALSE?

Think about it this example: suppose Y.H.P. promised you that

> If you study for the exam, then you will pass the course.

Now suppose you didn't study for the exam. Were any lies told, any promises broken? No, because Y.H.P. didn't say anything about what would happen if you *didn't* study for the exam—the claim was only about what you could expect if you *did* study for the exam. Accordingly, the conditional is TRUE in the case that the United States is not a democracy, because the conditional is silent on what happens in that case.

We can see this in the truth table for conditional, given in Table 4. Notice that the conditional is TRUE in the case that the antecedent φ is FALSE, regardless of the truth value of the consequent. We refer to statements like this as *vacuously true.*[33] But the more important pair of rows is the third and fourth rows, where the antecedent is TRUE. In that case, the conditional is FALSE in the case that the consequent is FALSE, and TRUE in the case that the consequent is TRUE—that's the bad prediction/good prediction pair of cases.

When given a conditional like $\varphi \Rightarrow \psi$, we say that φ is a *sufficient condition for* ψ and that ψ is a *necessary condition for* φ. In our example, we might say that being a democracy is a sufficient condition for holding elections, and that holding elections is a necessary condition for being a democracy. We can see this in the truth table: if φ is TRUE, then ψ must be TRUE in order for the conditional to be TRUE. That is, ψ is a necessary condition for φ. Similarly, if ψ is FALSE, then φ must be FALSE for the conditional to be TRUE. So, φ is a sufficient condition for ψ.[34]

[32] This is what Y.H.P. means by "hierarchical conditions." The idea is that we can put statements into some kind of hierarchy, where some statements are more fundamental than others. Here we have a statement about the United States being a democracy, which is more fundamental than a statement about the United States holding elections. We can use the more fundamental statement to make a prediction about the less fundamental statement via a conditional.

φ	ψ	$\varphi \Rightarrow \psi$
FALSE	FALSE	TRUE
FALSE	TRUE	TRUE
TRUE	FALSE	FALSE
TRUE	TRUE	TRUE

Table 4: Truth table for conditional.

[33] This is a fancy way of saying that the conditional is TRUE in the case that the antecedent is FALSE. It's a useful concept, and you'll be surprised how often it comes up.

[34] Of course, you're free to stop and think about whether this is all true in the real world: is it really the case that being a democracy is a sufficient condition for holding elections? Y.H.P. thinks so, but he's been wrong before. And surely the *converse* is not true: holding elections is not a sufficient condition for being a democracy. Many authoritarian regimes hold elections, but they are not democracies. This then raises the fun question: what is democracy that goes beyond just holding elections? What is the "extra" part that makes us feel like there's something missing in characterizing democracy only by the presence of elections? Perhaps it has to do with the protection of individual rights, the rule of law, or the presence of checks and balances on power. Or that other thing I told you about.

NEXT UP IS BICONDITIONAL. This is also important, because it tells us what it means for two statements to be logically equivalent. Consider:

> The United States is a democracy if and only if its leaders are chosen via free and fair elections.[35]

Let $\mathbf{D}(s)$ mean "s is a democracy" and $\mathbf{E}(s)$ mean "s's leaders are chosen via free and fair elections." Then we can write

$$\mathbf{D}(\text{United States}) \iff \mathbf{E}(\text{United States}),$$

which encodes the statement that the United States is a democracy if and only if its leaders are chosen via free and fair elections. But again, what about the atoms of this molecule must be TRUE in order for the molecule to be TRUE? It turns out that we can split this molecule into two conditionals, one in each direction:

$$\mathbf{D}(\text{United States}) \implies \mathbf{E}(\text{United States}), \text{ and}$$
$$\mathbf{E}(\text{United States}) \implies \mathbf{D}(\text{United States}).$$

The first of these says that if the United States is a democracy, then its leaders are chosen via free and fair elections. The second of these says that if the United States's leaders are chosen via free and fair elections, then the United States is a democracy.[36] In other words, democracy is both necessary *and* sufficient for leaders being chosen via free and fair elections, and leaders being chosen via free and fair elections is both necessary *and* sufficient for democracy. This is what we mean by logical equivalence.

We can see this in the truth table for biconditional, given in Table 5. Here Y.H.P. has added two new columns, one for each of the two conditionals that make up the biconditional. Notice that the biconditional $\varphi \Leftrightarrow \psi$ is just the "and" of the two conditionals $\varphi \Rightarrow \psi$ and $\psi \Rightarrow \varphi$. Notice also that this means that $\varphi \Leftrightarrow \psi$ is TRUE in the case that $\varphi \Rightarrow \psi$ is TRUE and $\psi \Rightarrow \varphi$ is TRUE, and FALSE otherwise—in other words, when φ and ψ have the same exact truth value, either both TRUE or both FALSE.

ONE OFTEN USES BICONDITIONALS FOR DEFINITIONS. For example, suppose we wanted to *define* what makes a democracy a democracy. We might say something like this:

> A state is a democracy if and only if it holds free and fair elections.[37]

This means two things are going on at once:

1. If you knew that a state was a democracy, then you could predict that it holds free and fair elections.

2. If you knew that a state holds free and fair elections, then you could rightfully classify it as a democracy, using this definition.

So, you can go from abstraction to the real world, or you can use the real world to inform your abstractions.[38]

[35] Notice that we've walked "holding elections" back to "leaders are chosen via free and fair elections." One of these implies the other, but not vice versa—though all free and fair elections for leaders are held, not all elections that are held are free and fair. And of course, we're on the hook to say just what it is that we mean by "free and fair elections." This is important, because the truth of our biconditional depends on the precise meaning of these terms.

[36] How are you feeling about these two conditions, now that we've set a scope for "free and fair elections?" Y.H.P. is feeling pretty good about them. But then again, he doesn't study this stuff.

φ	ψ	$\varphi \Rightarrow \psi$	$\psi \Rightarrow \varphi$	$\varphi \Leftrightarrow \psi$
FALSE	FALSE	TRUE	TRUE	TRUE
FALSE	TRUE	TRUE	FALSE	FALSE
TRUE	FALSE	FALSE	TRUE	FALSE
TRUE	TRUE	TRUE	TRUE	TRUE

Table 5: Truth table for biconditional.

[37] They tell Y.H.P. that he ought to use the same examples again and again to make it easier for you to follow along. But man, he's tired about figuring out whether boujee North American states are democracies.

[38] Whoa, we just bumped into the relationship between the real world and our theories of it. You'll see that this class will force you to think about this relationship a lot.

So we've learned five ways to link statements together: negation ($\neg$), conjunction ($\wedge$), disjunction ($\vee$), conditional ($\Rightarrow$), and biconditional ($\Leftrightarrow$)—informally, you can read these as "not," "and," "or," "if-then," and "if-and-only-if." You'd be surprised just how many claims about I.R. one can encode using just those five connectives. But, sometimes we want to make more general claims, and that's where *quantifiers* come in.[39]

Suppose that Y.H.P. thought his claim about democracies and elections wasn't just true for the United States, but for all democracies. One might write something like

All democracies hold elections.

Let me re-write this into a form that makes the quantifier more obvious:

For all states s, if s is a democracy, then s holds elections.[40]

This is the kind of statement that we can encode using a *universal quantifier,* which is a logical device that allows us to make claims about all objects in the universe of discourse. The symbol is $\forall$, and you can read it as "for all." Thus, the formal statement is

$$\forall s\,(\mathbf{D}(s) \Longrightarrow \mathbf{E}(s)),$$

which is read as "for all states s, if s is a democracy, then s holds elections."[41] That seems a faithful representation of the claim that all democracies hold elections, right?

Now, this is a more complicated molecule than we've seen before, for two reasons:

1. we're making a big chained molecule out of a bunch of small molecules, one for each state in the universe of discourse; and
2. each of those molecules is a conditional, rather than an atomic statement itself.

Well, we can handle the second complication with a definition, say

$$\mathbf{P}(s) \Longleftrightarrow (\mathbf{D}(s) \Rightarrow \mathbf{E}(s)),$$

so that $\mathbf{P}(s)$ is TRUE just in case State s satifies $\mathbf{D}(s) \Rightarrow \mathbf{E}(s)$.[42] That's just to simplify our notational lives. Now that we've gotten that out of the way, we can handle the first complication with a big ol' conjunction:

$$\mathbf{P}(s_1) \wedge \cdots \wedge \mathbf{P}(s_n) \cdots$$

which is TRUE just in case all of the underlying $\mathbf{P}(s_i)$s are TRUE. That's what we mean by "for all states s."[43] This treatment might seem a bit heavy-handed, but it's actually quite useful: it allows us to talk about all of the states in the universe of discourse at once, rather than having to talk about them one at a time, or even handful-by-handful. Once you get used to it, it's a lot easier to write and read. What's more, it allows you to gloss over the details of the underlying statements, which is often a good thing. But remember: all those little atoms are still there!

[39] These are called "quantifiers" because they tell us something about how many objects in the universe of discourse satisfy some property. We can use such machines to talk about "all states," "no states," "at least one state," and so on.

[40] Notice that Y.H.P. has had no choice but to use the variable s here—one cannot talk about democracy *in the abstract* while also talking about a particular state. The study of democracy need not be the study of democracies. *Whoa.*

[41] Notice that the $\forall$ quantifier demands that we speak of the universe of discourse—here states—to be able to talk about the objects in it. In other words, to talk about something is to talk about *some thing,* or more specifically *some kind of thing.*

[42] Can you re-write $\mathbf{P}(s)$ using only $\neg$ and $\wedge$? *Hint:* notice that $\mathbf{P}(s)$ is TRUE when either $\mathbf{D}(s)$ is FALSE or $\mathbf{E}(s)$ is TRUE. *Hint for the hint:* try to prove to yourself with a truth table that

$$(\varphi \Longrightarrow \psi) \Longleftrightarrow (\neg\varphi \vee \psi).$$

[43] Notice that Y.H.P. has left open the possibility that there are infinite states in the universe of discourse—that's why the expression ends in "$\cdots$." It could be that there are only a finite number of states, in which case the expression would end in "$\wedge \mathbf{P}(s_n)$," where n is the number of states in the universe of discourse. So, in the real world we're currently experiencing, there is in fact a finite number of states, and it's in the ballpark of two hundred or so. But the universe of all *possible* states is infinite!

One more to go. Suppose Y.H.P. wanted to say that at least one non-democracy holds elections.[44] We might write something like

> Some state that is not a democracy holds elections.

Let me re-write this into a form that makes the quantifier more obvious:

> There exists a state s such that s is not a democracy and s holds elections.

To encode this, we use an *existential quantifier,* which is a logical device that allows us to make claims about at least one object in the universe of discourse. The symbol is $\exists$, and you can read it as "there exists."[45] Thus, the formal statement is

$$\exists s \left(\neg \mathbf{D}\left(s\right) \wedge \mathbf{E}\left(s\right)\right),$$

which is read as "there exists a state s such that s is not a democracy and s holds elections."

Whereas the universal quantifer $\forall$ is like a big conjunction—that is, a big "and"—the existential quantifier $\exists$ is like a big disjunction—that is, a big "or." Again to simplify notation, define

$$\mathbf{Q}\left(s\right) \iff \left(\neg \mathbf{D}\left(s\right) \wedge \mathbf{E}\left(s\right)\right),$$

which is just the statement that s is not a democracy and s holds elections. Then we can write

$$\mathbf{Q}\left(s_1\right) \vee \cdots \vee \mathbf{Q}\left(s_n\right) \cdots$$

which is `TRUE` just in case at least one of the underlying $\mathbf{Q}\left(s_i\right)$s is `TRUE`.[46]

Here's a useful exercise that will help you identify a common mistake. Consider again the statement

> All democracies hold elections.

Now, suppose you thought this statement was `FALSE`; how would you corroborate your instinct? What kind of evidence would you need to show that all democracies do not hold elections? Would you need to show that all democracies *don't* hold elections? Surely not, right? All you'd need to do is show that *one* democracy doesn't hold elections.[47] In other words, to show that a "for all" statement is `FALSE`, you only need to show that *one* of the underlying statements is `FALSE`, which in turn requires a "there exists" statement. In other words,

$$\neg \forall s \left(\mathbf{D}\left(s\right) \implies \mathbf{E}\left(s\right)\right) \iff \exists s \left(\neg \mathbf{D}\left(s\right) \wedge \mathbf{E}\left(s\right)\right),$$

where the left-hand side reads "it is not the case that for all states s, if s is a democracy, then s holds elections," and the right-hand side reads "there exists a state s such that s is not a democracy and s holds elections." A similar exercise can be done for "there exists" statements: to show that a "there exists" statement is `FALSE`, you only need to show that *none* of the underlying statements is `TRUE`, which in turn requires a "for all" statement.[48]

[44] Remember, Adolf Hitler was elected, though one wouldn't call the 1933 German election "free and fair:" it took place a few weeks after the Nazi Party finished seizing power, and those weeks included plenty of violence and intimidation.

[45] For the record, we won't be using this level of formalism all the time. But, you'd be surprised how much better you get at using plain language precisely from trying to encode it in formal language. And precision, student of mine, is the name of the game. Don't confuse precision with rigidness, though—you can be precise and whimsical (or creative, or playful, or...) at the same time. Indeed, Y.H.P. would argue that the whimsy is more whimsical when it's precise.

[46] Consider what kind of evidence it would take to corroborate a "for all" statement versus an "there exists" statement. For example, to corroborate the statement that all democracies hold elections, you'd have to show that *each and every democracy* holds elections. However, to corroborate the statement that at least one non-democracy holds elections, you'd only have to show that *one non-democracy* holds elections. It doesn't matter if you can identify one, or two, or three, or a hundred non-democracies that hold elections—all that matters is that you can identify *at least one.*

[47] We call such a democracy a *counterexample* to the statement that all democracies hold elections. You should try to develop comfort with the idea of counterexamples, because they're a really useful tool for thinking about the truth of general claims.

[48] In other words,

$$\neg \exists s \left(\mathbf{R}(s)\right) \iff \forall s \left(\neg \mathbf{R}(s)\right),$$

where $\mathbf{R}(s)$ is any statement you'd like about State s.

You'd be surprised just how many people on cable news struggle with this. Actually, if you thought about it, you wouldn't be surprised; those people are nearly uniformly stupid.

THAT WAS A LOT, so let's try an extended example to put it all together. Suppose we wanted to say that all democracies hold elections, but that some non-democracies also hold elections. Again, let $\mathbf{D}(s)$ encode "s is a democracy" and $\mathbf{E}(s)$ encode "s holds elections." The first part of the claim—that all democracies hold elections—is

$$\forall s\, (\mathbf{D}(s) \implies \mathbf{E}(s)),$$

and the second part of the claim—that some non-democracies also hold elections—is

$$\exists s^*\, (\neg \mathbf{D}(s^*) \wedge \mathbf{E}(s^*)),$$

where we use the special symbol s^* to denote that we're talking about a *particular* state, rather than all states.[49] Now, we can put these two statements together using a conjunction:

$$(\forall s\, (\mathbf{D}(s) \implies \mathbf{E}(s))) \wedge (\exists s^*\, (\neg \mathbf{D}(s^*) \wedge \mathbf{E}(s^*))),$$

which faithfully encodes the claim.[50]

[49] This is just a convention, though—we could use any symbol we wanted, such as t, s', whatever.

[50] And *that's* how you waste an hour talking about how to write down a fact one learns in PS 100. Can this man teach, or what? This man is a *professional.*

IF YOU CAN BELIEVE IT, all of this nonsense raises a really important question for I.R.: what is the universe of discourse?[51] In particular, when we say "for all" or "there exists a," we need to be clear about what's covered by the "all" or "a." For example, if the "for all" covers, like, *everything,* then "For all s, s is a democracy" means that we have

Y.H.P. is a democracy.

[51] The philosophical study of such questions is called *ontology,* and it's a really important part of I.R.—and yet it's often ignored. You know that a theory is *bad* when it doesn't even bother to tell you what it's talking about.

Now look: Y.H.P. is a lot of things, but he's not a democracy; we've again landed on one of those "is this well-formed?" questions. But these are ubiquitous in I.R., especially when you think about all the levels of analysis we use:

1. states;
2. individuals, be they leaders or citizens;
3. groups, be they lobbying firms, political parties, or social movements;
4. alliances;
5. commodities, be they oil, gold, or bananas;
6. international organizations, be they the United Nations, the World Trade Organization, or the International Monetary Fund.

But there are also different *kinds* of states, individuals, and so on, so we might have to think about the universe of discourse in terms of *sorts* of objects, rather than just objects. And, these objects interact with various modalities, so that we might need to say "for all wars" or "there exists a sanction." In other words, I.R. is an absolute ontological mess, and we need to be careful about what we're talking about when we talk about it.

So, the logic Y.H.P. has been showing you is called *first-order logic.* It allows us to quantify over things using ∀ and ∃, but it doesn't allow us to quantify over different *sorts* of things. Y.H.P. has been sloppy by saying "For all *states s*," because that implicitly sets the one allowable sort of thing to be states. But if Y.H.P. wanted to say the (altogether reasonable) statement

For all states s there exists an individual i such that i is the leader of s,

then he'd have to be more careful about the universe of discourse: now we're quantifying over both states and individuals. This takes us from traditional first-order logic to something called *many-sorted first-order logic,* which isn't much more complicated than what we've been doing, but it's a bit more formal. Yet, it would seem that the unique challenges of I.R. inform our decisions not just about what to talk about (substance), but *how* to talk about it (which logic). Trippy, right? Methodology is a *big deal.*

THE GOAL OF THESE TOOLS is to help us make valid claims.

2 Definition (Validity)

A claim is valid *if and only if it is* TRUE *in all possible cases.*

Let's consider a well-known example: that of the *contrapositive.*

Contrapositives really help you to see the necessary/sufficient thing. When we write $\varphi \Rightarrow \psi$, we're saying that φ is a sufficient condition for ψ—so, if a state is a democracy, you know that it holds elections. But, the contrapositive $\neg\psi \Rightarrow \neg\varphi$ says that $\neg\psi$ is a sufficient condition for $\neg\varphi$—so, if a state doesn't hold elections, you know that it's not a democracy.

3 Definition (Contrapositive)

The contrapositive *of a conditional* $\varphi \Rightarrow \psi$ *is the conditional* $\neg\psi \Rightarrow \neg\varphi$.

It turns out that a conditional and its contrapositive are logically equivalent: one of them is TRUE if and only if the other is TRUE. Let's show that with a truth table. First things first, let's write out all four cases, including both the "raw" statements and their negations:

Raw		*Negated*	
φ	ψ	$\neg\varphi$	$\neg\psi$
FALSE	FALSE	TRUE	TRUE
FALSE	TRUE	TRUE	FALSE
TRUE	FALSE	FALSE	TRUE
TRUE	TRUE	FALSE	FALSE

Remember: if there are n statements being combined, then there are 2^n cases. So here, we have two raw statements—φ and ψ—and thus four cases. Columns 3 and 4 are obtained by using the rules of negation.

Now, let's use φ and ψ to add a column for $\varphi \Rightarrow \psi$:

Raw		*Negated*		*Conditional*
φ	ψ	$\neg\varphi$	$\neg\psi$	$\varphi \Rightarrow \psi$
FALSE	FALSE	TRUE	TRUE	TRUE
FALSE	TRUE	TRUE	FALSE	TRUE
TRUE	FALSE	FALSE	TRUE	FALSE
TRUE	TRUE	FALSE	FALSE	TRUE

Notice that we've just used the truth table for conditional to fill in the fifth column. The first two rows are vacuously TRUE, because those are the cases where the antecedent φ is FALSE—no promises have been broken; Y.H.P. has taken the libery of coloring those cells orange. The third row is FALSE, because the antecedent φ is TRUE but the consequent ψ is FALSE—a promise has been broken. The fourth row is TRUE, because the antecedent φ is TRUE and the consequent ψ is TRUE—a promise has been fulfilled.

Finally, let's use $\neg\psi$ and $\neg\varphi$ to add a column for the contrapositive, $\neg\psi \Rightarrow \neg\varphi$:

Raw		*Negated*		*Conditional*	*Contrapositive*
φ	ψ	$\neg\varphi$	$\neg\psi$	$\varphi \Rightarrow \psi$	$\neg\psi \Rightarrow \neg\varphi$
FALSE	FALSE	TRUE	TRUE	TRUE	TRUE
FALSE	TRUE	TRUE	FALSE	TRUE	TRUE
TRUE	FALSE	FALSE	TRUE	FALSE	FALSE
TRUE	TRUE	FALSE	FALSE	TRUE	TRUE

Pay in mind that the final column uses $\neg\psi$ as the antecedent, so that it's "Column 4 implies Column 3," rather than "Column 3 implies Column 4." The vacuous cases should clarify how that's all working.

As you can see, the conditional and its contrapositive are logically equivalent: they're TRUE in the same cases, and FALSE in the same cases.[52]

[52] Such statements are called *tautologies,* and they're TRUE in all possible cases. Sometimes one hears "that's tautological!" in a pejorative sense, but that usually means something like "that's baked into the definition of the thing you're talking about!" or "there is no conceptual daylight between the two things you're talking about!" Tautologies are awesome, though, assuming that there is indeed conceptual daylight between the two things you're talking about.

LET'S DO ANOTHER! To do so, let me define an important concept: *transitivity*.[53] The following two conditional seem to be true:

1. If state s is a liberal democracy, then state s is a democracy.
2. If state s is a democracy, then state s holds elections.

Now, you might have drawn a conclusion, namely

If state s is a liberal democracy, then state s holds elections.

But you weren't given that conditional, so you can't just assume it's true. Define

$$\mathbf{L}(s) \rightsquigarrow \text{"}s \text{ is a liberal democracy"}$$
$$\mathbf{D}(s) \rightsquigarrow \text{"}s \text{ is a democracy"}$$
$$\mathbf{E}(s) \rightsquigarrow \text{"}s \text{ holds elections."}$$

And what we want is something like

If $\mathbf{L}(s)$ implies $\mathbf{D}(s)$, and $\mathbf{D}(s)$ implies $\mathbf{E}(s)$, then $\mathbf{L}(s)$ implies $\mathbf{E}(s)$,

which we can write formally as

$$[(\mathbf{L}(s) \Longrightarrow \mathbf{D}(s)) \wedge (\mathbf{D}(s) \Longrightarrow \mathbf{E}(s))] \Longrightarrow (\mathbf{L}(s) \Longrightarrow \mathbf{E}(s)).$$

This is a conditional, so we can use the truth table for conditional to figure out when it's TRUE and when it's FALSE. Since we have three atomic statements, we'll have eight cases. Let's gather all the materials we need:

Raw			*Derived*			
$\mathbf{L}(s)$	$\mathbf{D}(s)$	$\mathbf{E}(s)$	$\mathbf{L}(s) \Rightarrow \mathbf{D}(s)$	$\mathbf{D}(s) \Rightarrow \mathbf{E}(s)$	$(\mathbf{L}(s) \Rightarrow \mathbf{D}(s)) \wedge (\mathbf{D}(s) \Rightarrow \mathbf{E}(s))$	$\mathbf{L}(s) \Rightarrow \mathbf{E}(s)$
FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE
FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE
FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE
TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE
TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

Proof that the next-to-last column implies the last column as an exercise.[54]

Here's a general definition of transitivity.

4 Definition (Transitivity)

Given a universe of discourse and a relationship $\mathbf{R}$ *between objects in said universe, we say that* $\mathbf{R}$ *is* transitive *just in case, for any three objects* x, y, *and* z *in the universe, if* x *is* $\mathbf{R}$*-related to* y *and* y *is* $\mathbf{R}$*-related to* z, *then* x *is* $\mathbf{R}$*-related to* z. *Any relationship that is not transitive is* intransitive.[55]

Here the universe is statements, and the relationship is implication.

[53] Fair warning: transitivity is going to be your life for the next few weeks. You'd be surprised just how many things rely on it—like, deeply substantive things. For example, we'll define rationality in terms of transitivity. But think about any way that we put things into order: we're relying on transitivity. If you tried to rank all states by their wealth or power, wouldn't you wind up with a transitive ranking?

In general, when Y.H.P. uses the squiggly $\rightsquigarrow$, it's being used in a more informal sense than the $\Longrightarrow$ symbol. In the study of formal languages, we might use the term "points to" to describe the relationship between a symbol and its meaning.

[54] Informally, notice that the only time the last column is FALSE happens to be when the next-to-last column is also FALSE, so we always get saved by vacuous truth. Notice, however, that the claim does *not* work in reverse: the last column does *not* imply the next-to-last column.

[55] Notice the role that vacuous truth can play here: if there are no chains $x \longrightarrow y \longrightarrow z$, then there are no promises to be broken. And it's got to be the same y that serves as the middle term in the chain, otherwise we're talking about two unrelated chains $x \longrightarrow y_1$ and $y_2 \longrightarrow z$. We'll get a lot of practice with this in a few weeks, so just make a note of it for now.

But in the meantime, riddle me this: is the relationship $\geqslant$ transitive on the universe of discourse of the counting numbers? That is, if you had any three numbers n_1, n_2, and n_3, and you knew that $n_1 \geqslant n_2$ and $n_2 \geqslant n_3$, would you also know that $n_1 \geqslant n_3$?

TRANSITIVITY SEEMS PRETTY STRAIGHTFORWARD, and yet it is super powerful. If you know that a relationship between things is transitive, then you can make a lot of claims about the things in the universe of discourse. Just to make this easy, let the universe of discourse be individuals for now, and consider the statement

$\mathbf{A}\left(i_1, i_2\right) \rightsquigarrow$ "i_1 is an ancestor of i_2."

This is clearly transitive: if one's grandmother is an ancestor of one's mother, and if one's mother is an ancestor of oneself, then one's grandmother is an ancestor of oneself.[56] Now, you might be wondering to yourself: what about one's great-grandmother, who we know is an ancestor of one's grandmother? Do we know that one's great-grandmother is an ancestor of oneself, too? Well, yeah—but can you say why?

[56] *For this,* you're paying. But now consider this, chucklehead: letting $\mathbf{M}\left(i_1, i_2\right)$ convey "i_1 is a mother of i_2," is **M** transitive? Now suppose that one's mother was adopted, so that (biologically speaking only!) one's grandmother is not an ancestor of one's mother. Does this mean that we need one's grandmother to be an ancestor of oneself? Think about vacuous truth, broken promises, etc.

5 Proposition (Chained Transitivity)

Consider a universe with a transitive relation **R**. *For any n objects $x_1, x_2, \ldots, x_n$, if x_1 is* **R***-related to x_2, x_2 is* **R***-related to x_3, ..., and x_{n-1} is* **R***-related to x_n, then x_1 is* **R***-related to x_n.*

The idea behind Proposition 5 is straightforward. Suppose that we had four objects called x_1, x_2, x_3, and x_4, and suppose that x_1 is **R**-related to x_2, x_2 is **R**-related to x_3, and x_3 is **R**-related to x_4. In the figure below, this is represented by the three gray arrows. Now, since **R** is transitive, our assumptions that $x_1 \mathbf{R} x_2$ and $x_2 \mathbf{R} x_3$ imply that $x_1 \mathbf{R} x_3$. That means we can

This is the very essence of putting things in order. We'll be doing this often throughout the class; in rational choice, we proceed as if decision-makers can put things in order, and in game theory we proceed as if players can put strategies in order.

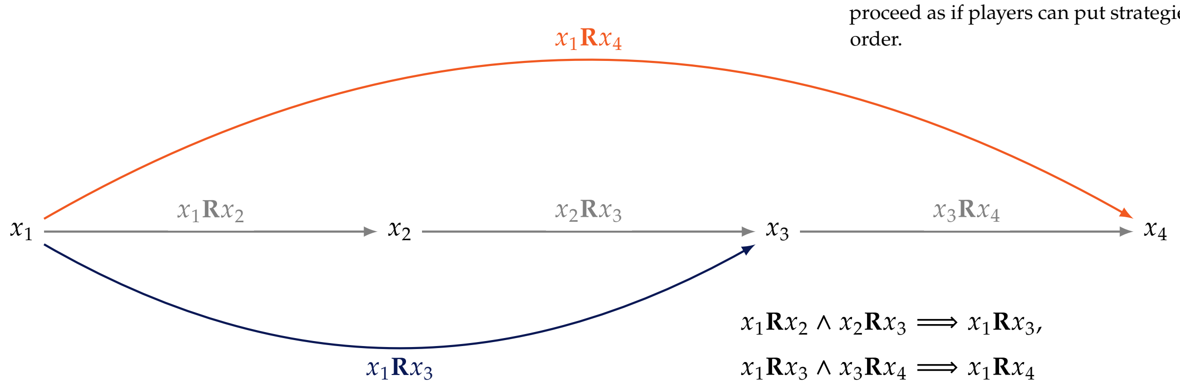


Figure 1: Chained transitivity: the case of $n = 4$.

proceed on the understanding that $x_1 \mathbf{R} x_3$. Well, this inference that $x_1 \mathbf{R} x_3$ and our assumption that $x_3 \mathbf{R} x_4$ imply that $x_1 \mathbf{R} x_4$. If there happens to exist some object x_5 such that $x_4 \mathbf{R} x_5$, then we can proceed on the understanding that $x_1 \mathbf{R} x_5$. And so on. For example, suppose you were thinking about the counting numbers and that $\mathbf{R} = \leqslant$. Then, if you knew that $1 \leqslant 2$ and $2 \leqslant 3$, you could infer that $1 \leqslant 3$. From there, you could infer that $1 \leqslant 4$, and so on. In fact, you could infer that $1 \leqslant n$ for any counting number n—and that's exactly what Proposition 5 says.

You prove the result in two steps:

1. *Base case:* prove that the claim holds for chains with $n = 3$ objects, which is the smallest possible chain.
2. *Induction step:* prove that if the claim holds for chains with n objects, it must hold for chains with $n + 1$ objects.

That's exactly the logic depicted in Figure 1 and described at left.

Sets

Now look, we're not going to go through this whole class speaking just in formal languages. Though logic is certainly part of mathematics—or at least is *highly* math-adjacent—it's not the whole of mathematics.[57] And we most definitely will not be speaking in terms of formal logic over these weeks we have together. We need a way to send the things we're talking about into mathematical structures we can work with, and that's where *sets* come in. Nearly all of mathematics is built on sets.

[57] You'd be surprised just how mathy logic can get—indeed, toward the turn of the 20th Century folks got the idea that all of mathematics could be reduced to logic.

Consider the counting numbers,[58]

$$1, 2, 3, 4, 5, \ldots.$$

[58] Some people consider 0 to be a counting number. Y.H.P. can't quite figure out how he feels about that.

You use these all the time, right? Counting is fun.[59] Well, what if we wanted to talk about the counting numbers as one thing, rather than a number-by-number list? We can *gather* the counting numbers into a *set*, which we'll call $\mathbb{N}$.[60] To demonstrate that we've done that, we use curly brace symbols, $\{\}$, and we write

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\},$$

[59] You might even think that counting is a uniquely human thing, but it turns out that other species can count, too. Among these are dolphins, elephants, and crows.

[60] The font used for $\mathbb{N}$ is called "blackboard bold," and it's used to denote common sets.

where the ellipsis "..." is just a shorthand for "and so on."[61] Now, we can talk about the counting numbers as a single thing, rather than a list of things—indeed, the father of set theory, Georg Cantor, defined a set through the poetic (but insufficiently precice) "out of many, one."[62]

Just in the name of completeness, here's a working definition of a set.

---

6 Definition (Sets)

A set *is a collection*[63] *of objects, called* elements.

The objects in a set can be anything:

1. numbers, like the counting numbers or the real numbers;
2. states, like the states of the world, or all the world's democracies;
3. actions, like wars or sanctions;
4. individuals, like the citizens of the United States or the board members of Amnesty International;
5. coalitions, like all possible bilateral alliances, all possible trilateral alliances, or all possible alliances of any size;
6. commodities, like oil, gold, or bananas; or
7. anything else you can think of.

Heck, the elements of a set can even be other sets! You'd be surprised how helpful that can be—consider the set of all possible global alliance structures, which (as you'll see on your problem set) is just a set of sets.

[61] The ellipsis is a *really* important symbol in mathematics, and it's used to denote that something is continuing in a pattern. For example, if you wanted to write out the counting numbers, you could write

$$1, 2, 3, 4, 5, \ldots, 100,$$

which would be understood to mean "the counting numbers from 1 to 100." But if you wanted to write out the counting numbers *forever,* you could write

$$1, 2, 3, 4, 5, \ldots,$$

which would be understood to mean "the counting numbers from 1 to ∞." The ellipsis is also used to denote that something is continuing in a pattern, but that the pattern is *not* obvious. For example, if you wanted to write out the counting numbers, but you wanted to leave out the even numbers, you could write

$$1, 3, 5, 7, 9, \ldots,$$

which would be understood to mean "the counting numbers from 1 to ∞, but leave out the even numbers."

[62] The Latin phrase *e pluribus unum* is the motto of the United States, and it means "out of many, one." It's a reference to the fact that the United States is a union of many states, but it's also a reference to the fact that the United States is a union of many people. The idea is that the United States is a single thing, but it's made up of many things. The same is true of sets: a set is a single thing, but it's made up of many things.

[63] Naturally, we have defined the term "set" by way of the undefined term "collection." This is a bit of a problem, but it's not disastrous.

Elementhood is the key: sets represent nothing but a collection of objects, so the only existential claim we can make about a set is that an object is an element of the set (or that it is not).

---

7 Definition (Elementhood)

Let X be a set. If some object x is included in X, then we say that x is an element *of X, and we write $x \in X$. If x is not included in X, then we write $x \notin X$.*

So for example, consider the set of card suits,[64]

$$S = \{♥, ♣, ♠, ♦\}.$$

This gives us four statements[65] about elementhood:

♥ $\in S$ $\rightsquigarrow$ "♥ is an element of S."

♣ $\in S$ $\rightsquigarrow$ "♣ is an element of S."

♠ $\in S$ $\rightsquigarrow$ "♠ is an element of S."

♦ $\in S$ $\rightsquigarrow$ "♦ is an element of S."

But we also have an infinite number of statements about non-elementhood; for example, observe that

$1 \notin S$ $\rightsquigarrow$ "1 is not an element of S."

$2 \notin S$ $\rightsquigarrow$ "2 is not an element of S."

$3 \notin S$ $\rightsquigarrow$ "3 is not an element of S."

$\vdots$

And so on, forever.[66]

Indeed, elementhood is so vital to sets that it's the only thing we use to define them. In particular:

1. The order of elements doesn't matter, so that (for example) $\{1,2,3\} = \{1,3,2\} = \{2,1,3\} = \{2,3,1\} = \{3,1,2\} = \{3,2,1\}$.[67]
2. The number of times an element is listed doesn't matter, so that (for example) $\{1,1,2,3\} = \{1,2,3\}$.

In other words, sets are *unordered* and *unique*. We might care about the order of things, and we might care about the number of times things are listed, but sets don't care about those things. We'd have to use a different sort of theoretical container to store objects in ways that care about order and/or uniqueness.[68] And we will do so, actually—just not quite yet. And really, we don't need order or uniqueness for the most foundational things one does in i.r. theory, which is stating and organizing the objects of interest, be they states, individuals, coalitions, or whatever else. In other words, sets help us to appreciate a theory's underlying ontology: the things that it takes to be real.

Throughout, we'll use capital letters to denote sets.

Symbol	Name
♥	Hearts
♣	Clubs
♠	Spades
♦	Diamonds

Table 6: Card suits.

[64] y.h.p. is struck by how few people know the names of the card suits. Please see Table 6 for a refresher.

[65] Notice that these are indeed statements: they're either TRUE or FALSE, not both, and not neither. The statement ♥ $\in S$ is TRUE, because ♥ is an element of S. The statement ♥ $\notin S$ is FALSE, because ♥ is an element of S. The statement ✠ $\in S$ is FALSE, because ✠ is not an element of S. The statement ✠ $\notin S$ is TRUE, because ✠ is not an element of S.
How about these?

$\neg$(♥ $\in S$)?

♥ $\in S \land$ ✠ $\in S$?

♥ $\in S \lor$ ✠ $\in S$?

♥ $\in S \Rightarrow$ ✠ $\in S$?

✠ $\in S \Rightarrow$ ♥ $\in S$?

[66] And that's just counting numbers! We could also say that "The United States is not an element of S." or whatever else you can think of that's not a card suit.

[67] By using the equals sign here, we're saying that these sets are *identical*. The word "identical" means that they have the same identity: we ignore differences in order, assigning all six of the sequences of symbols just listed "the same identity." Consider the breezy way that you write things like $1/2 = 2/4 = 3/6 = 4/8 = \ldots$. Well, are those really the same thing? Yes, they are, because they have the same identity, but that's because we've agreed to identify fractions by their decimal expansions.

[68] Can you think of something where the order of elements matters? Well, consider a *process*—say, the process by which two states negotiate a bilateral free trade agreement. The order in which the states make their demands matters, because the first state to make a demand might be able to extract more concessions from the second state. A process is simply a sequence of steps that are taken, and the order in which those steps are taken matters. You typically don't make a pie by filling the crust with apples and then rolling out the dough.

SINCE WE USE SETS to gather things we care about, we should talk about how to introduce them. We've already seen one way to do that: we can list the elements of a set, which is called *roster notation.* We've done this a few times:

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\},$$
$$S = \{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\}.$$

Notice that roster notation can allow for infinite sets, as with $\mathbb{N}$, or finite sets, as with S.[69] We can even allow the ellipsis on both sides! Consider the complete set of integers,

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

This is every whole number, tending down to $-\infty$ and up to ∞.[70]

BUT ROSTER NOTATION is not the only way to introduce a set. We can also use *set-builder notation,* which is a bit more abstract. Instead of writing out the elements of a set, we write a *predicate* that describes the elements of the set.[71] For a numerical example, let's try to write out the set of all even numbers.[72] What makes an even number an even number? Well, it's a number that's divisible by 2, and we'd write this out as

$$E = \{x \in \mathbb{Z} \mid x \text{ is divisible by } 2\}.$$

So what's going on here? Four things:

1. We're defining a set called E.
2. Inside the curly braces, we use the $x \in \mathbb{Z}$ to introduce a variable x that ranges over the integers. That sets a scope limitation for the set: we will be defining it in terms of properties of integers.
3. The vertical bar "|" is read "such that," and it's used to separate the variable from the predicate.
4. To the right of the vertical bar, we write out the predicate that defines the set. In this case, we say that x is divisible by 2; to be included in the set, this must be TRUE.

So, we read the definition above as follows:

> The set E is the set of all integers x such that x is divisible by 2.

Thus, we have $2 \in E$, since 2 is divisible by 2. We also have $1 \notin E$, since 1 is not divisible by 2. And so on.

It should be noted that one sometimes lazily introduces a set just by writing out something that suggests the predicate, as in

$$\{\textit{ even numbers }\}, \text{ or}$$
$$\{\textit{ democratic states }\}, \text{ or}$$
$$\{\textit{ coalitions of size } 3\}.$$

[69] We say a set is *finite* if it has a finite number of elements, and we say a set is *infinite* if it has an infinite number of elements. Shocking, right?

[70] What is ∞? What is $-\infty$? Oh jeez, don't even get Y.H.P. started. For our purposes, it's useful to keep things simple: ∞ is a symbol we use to denote "a quantity larger than the quantity encoded by any number." Similarly, $-\infty$ is a symbol we use to denote "a quantity smaller than the quantity encoded by any number."

[71] Remember that a predicate is just something that allows us to make a statement about an object's properties. We had

> – is a democracy.

as a predicate, which we used to make statements about states. It's the predicate of a descriptive sentence!

[72] Yes, we could do this in roster notation, as in $\{\ldots, -4, -2, 0, 2, 4, \ldots\}$, but that's not the point. The point is to introduce set-builder notation.

If we wanted to be a bit more precise, we could formalize the predicate as follows:

$$E = \{x \in \mathbb{Z} \mid \exists y \in \mathbb{Z}\,(x = 2y)\},$$

which reads "E is the set of all integers x such that there exists an integer y such that $x = 2y$."

More often than not, this is perfectly acceptable, and Y.H.P. will usually speak in these terms out loud.

So let's write out some special sets. You've already seen two very important sets, namely the counting numbers[73] and integers:

$$\mathbb{N} = \{1, 2, 3, 4, 5, \dots\},$$
$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

[73] For the record, one often calls $\mathbb{N}$ the *natural numbers.*

We can also define the set of *rational numbers,* which are numbers that can be expressed as a ratio of two integers:

$$\mathbb{Q} = \left\{ \frac{p}{q} \;\middle|\; p \in \mathbb{Z} \wedge q \in \mathbb{Z} \wedge q \neq 0 \right\}.$$

Let's read this definition aloud together:

> The set $\mathbb{Q}$ is the set of all numbers x such that there exist integers p and q such that $x = p/q$ and $q \neq 0$.

So here we have two variables, p and q, that range over the integers.

The rational numbers are all the numbers that you can write as a fraction, like $1/2$, $3/4$, $5/6$, and so on. But there are numbers that you can't write as a fraction, like π[74] or $\sqrt{2}$[75] or e.[76] In other words, $\pi \notin \mathbb{Q}$, $\sqrt{2} \notin \mathbb{Q}$, and $e \notin \mathbb{Q}$. But these are important numbers! We need a set that contains them! So we define the *real numbers,* which is the set of all numbers that fall strictly between $-\infty$ and ∞, be they rational or irrational:

$$\mathbb{R} = \{x \mid x \text{ is a number and } -\infty < x < \infty\}.$$

[74] π is the ratio of a circle's circumference to its diameter. It's an irrational number, which means that it can't be written as a fraction. As you might already know, $\pi \approx 3.14159$.

[75] $\sqrt{2}$ is the square root of 2, which is the number that you multiply by itself to get 2. It's also an irrational number. As you might already know, $\sqrt{2} \approx 1.41421$.

[76] e is the base of the natural logarithm, and it's also an irrational number. As you might already know, $e \approx 2.71828$.

This is an incredibly imprecise definition, but it's good enough for our purposes.[77] The real numbers are all the numbers that you can think of, be they rational or irrational. So we have $\pi \in \mathbb{R}$, $\sqrt{2} \in \mathbb{R}$, and $e \in \mathbb{R}$. However, it remains the case that $-\infty \notin \mathbb{R}$ and $\infty \notin \mathbb{R}$, because they're not numbers. They're just symbols we use to denote "a quantity larger than the quantity encoded by any number" and "a quantity smaller than the quantity encoded by any number," respectively.

[77] The issue is that the real numbers can't be stated *a priori* the way that the rational numbers can. The usual constructions of the real numbers are far too advanced for this class, so let's just go with the quick-and-dirty definition above.

We will use the real numbers all the time. Like, all the time all the time. It helps to have some good intuitions about them, and one of the best ways to do that is to think about the real numbers as a line. The real number line is just a line that extends from $-\infty$ to ∞, and it's got a tick mark for every real number. It's typical to put 0 in the middle, but that's not necessary—sometimes we care about 0, and sometimes we don't. In this example, we've simply ticked off some integers to give a sense of scale. However, you should remember that every real number falls along this line, not just the handful of integers that we've ticked off.[78] We sometimes care about the nonnegative reals or the strictly-positive reals,

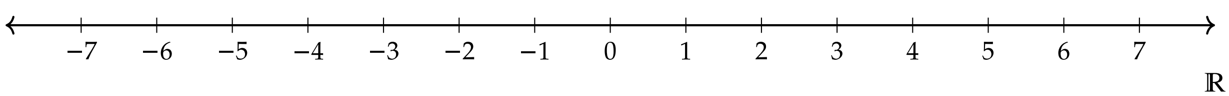


Figure 2: The real number line.

$$\mathbb{R}_{\geqslant 0} = \{x \in \mathbb{R} \mid x \geqslant 0\},$$
$$\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}.$$

Naturally, one can also define the nonpositive reals and the strictly-negative reals, but we won't use them much.

[78] Let's get trippy: the real numbers have an interesting property called *density.* This means that between any two real numbers, there is another real number; this, in turn, implies that there are an infinite number of real numbers between any two real numbers. For example, there are an infinite number of real numbers between 0 and 1, and there are an infinite number of real numbers between 0 and 2. Now consider: is one of those infinities larger than the other?

We also care about *intervals* of the real numbers: sets of real numbers that fall between two numbers. We have special notations for these.

Notation	Name	Definition
$[a, b]$	Closed interval	$\{x \in \mathbb{R} \mid a \leq x \leq b\}$
(a, b)	Open interval	$\{x \in \mathbb{R} \mid a < x < b\}$
$[a, b)$	Half-open interval	$\{x \in \mathbb{R} \mid a \leq x < b\}$
$(a, b]$	Half-open interval	$\{x \in \mathbb{R} \mid a < x \leq b\}$
$[a, \infty)$	Half-open interval	$\{x \in \mathbb{R} \mid a \leq x\}$
$(-\infty, b]$	Half-open interval	$\{x \in \mathbb{R} \mid x \leq b\}$
$(-\infty, \infty)$	Open interval	$\mathbb{R}$

So for (the most important) example, let $a = 0$ and $b = 1$, yielding

$$[0,1] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\},$$
$$(0,1) = \{x \in \mathbb{R} \mid 0 < x < 1\},$$
$$[0,1) = \{x \in \mathbb{R} \mid 0 \leq x < 1\},$$
$$(0,1] = \{x \in \mathbb{R} \mid 0 < x \leq 1\}.$$

We refer to the cases with $a = 0$ and $b = 1$ as "the unit interval."

Table 7: Intervals of the real numbers.

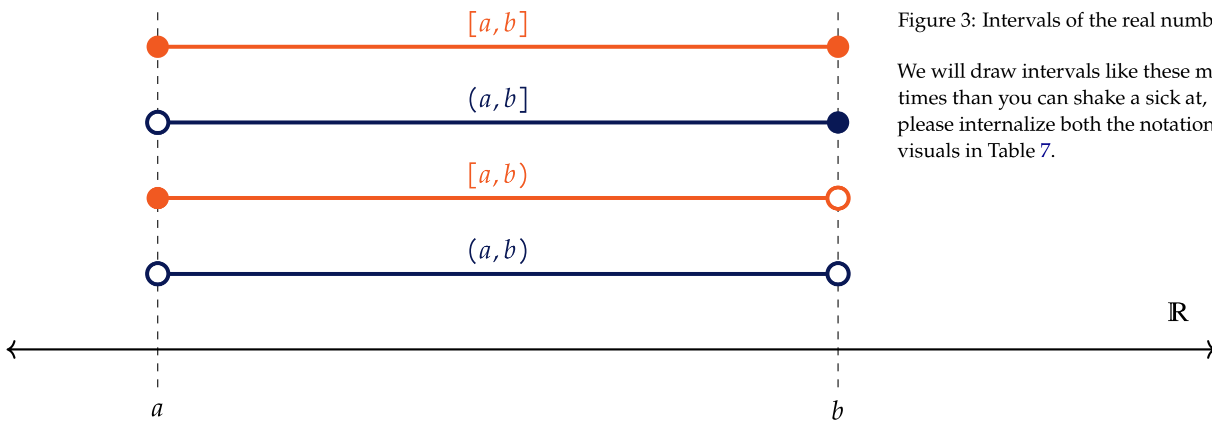


Figure 3: Intervals of the real numbers.

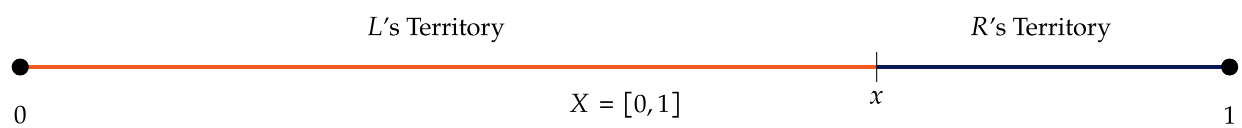
We will draw intervals like these more times than you can shake a sick at, so please internalize both the notation and visuals in Table 7.

We can use the real numbers to model many things in i.r.:

1. Suppose two states, call them L and R, were deciding where to draw a border. We can envision territory as a line,[79] and we can model the border as a point on that line. Formally, let $X = [0,1]$ denote all possible borders between L and R, where 0 is the leftmost point and 1 is the rightmost point.[80] Then, we can model the border as a point $x \in X$.

[79] Hello, strong assumption! We're assuming that territory is one-dimensional! Can you think of workarounds? We'll discuss some relevant tools pretty soon.

[80] Thus, if $x = 0$, R gets all the territory and L gets nothing. Conversely, if $x = 1$, L gets all the territory and R gets nothing.



We now can envision an infinite number of territorial situations. Yep, we just theorized an uncountably infinite number of borders—thanks, variables! y.h.p. thinks that's pretty cool, and that's just $[0,1]$.[81] Just put your finger where x is and slide it around![82]

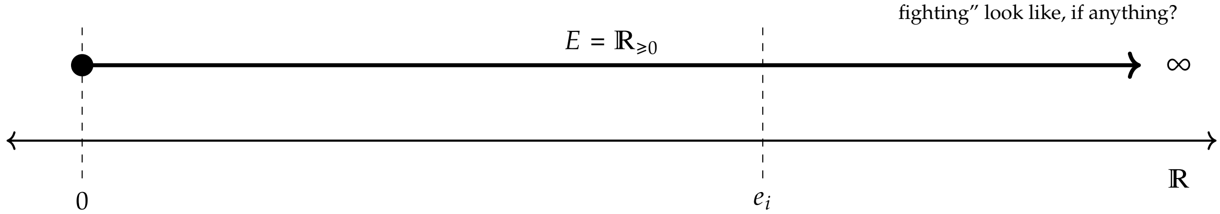
[81] What else could you model with $[0,1]$? Think about proportions, percentages, and so on. If there's an "all" and a "none" in the problem, you can probably model it with $[0,1]$.

[82] As you do this, imagine the screams of the people who live in the territory that you're taking away, the families you've separated, the lives you've ruined.

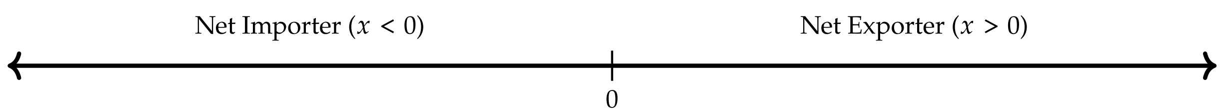
2. Suppose some state, call it i, found itself in a conflict. It must decide how many of its precious resources to mobilize for the war effort.[83] Let's suppose that there's a single resource and that State i can mobilize any non-negative amount of it; call the set of all available investments $E = \mathbb{R}_{\geqslant 0}$.[84] Call the amount of effort that they choose $e_i \in E$, which simply means $e_i \geqslant 0$ is a real number alerting us to the amount of resources that i has invested in the war effort.

[83] This is the beginning of the so-called *contest model of war,* which we'll discuss later on in the course. It is among Y.H.P.'s favorite models of war, because it's so simple and yet so powerful.

[84] Do you agree that the amount of resources invested in a war effort should be non-negative? What would "negative fighting" look like, if anything?



3. The *trade balance* for a state s on some commodity c is the difference between the amount of c that s exports and the amount of c that s imports. In case the trade balance is negative, we say that s has a *trade deficit* on c, or that it is a *net importer* of c. In case the trade balance is positive, we say that s has a *trade surplus* on c, or that it is a *net exporter* of c.[85] Let's denote the trade balance for state s on commodity c as x_s^c. Then, we can model the trade balance as a real number, $x_s^c \in \mathbb{R}$. By



allowing x_s^c to range over the (unbounded) real numbers, we're making a theoretical commitment, namely that the trade balance can be any real number. There is no good theoretical reason to cap the trade balance at some number, so we don't.[86]

[85] For example, the United States both imports and exports oil, but it imports more than it exports, so it has a trade deficit on oil.

[86] Contrast this with the first example: we have a good theoretical reason to cap the border at 0 and 1, because these represent "all" and "none," respectively. "All" and "none" are concepts that seemed relevant for the substantive problem at hand. Credits and debits are different: there's no good reason to cap them at some number. These abstract objects don't come from nowhere! They're part of the theory!

So, we can use the real numbers to model a lot of things in I.R.

ONE OF MY FAVORITE SETS is a bit of a weird one: the *empty set*. It contains no elements. Formally, we write

$$\varnothing = \{ \ \} .$$

Y.H.P. likes to call sets "containers" in part because it helps provide intuitions about the empty set.[87] The empty set is a container that contains nothing—an empty box.[88] Observe that for any object in the entire universe x, we know that $x \notin \varnothing$. You're not in the empty set, and neither is your dog, and neither is the United States, and neither is the President, nor the President's dog, nor the President's dog's fleas.[89] Nothing is in the empty set, because the empty set contains nothing.

[87] For the record, there is only one empty set. It's not like there are multiple empty sets, each of which is empty in its own way. There's only one way to be empty, at least with respect to elementhood.

[88] Contrast this with $\{\varnothing\}$: a box that contains an empty box. But then what is

$$\{\{\varnothing\}\}?$$

How about

$$\{\varnothing, \{\varnothing\}\}?$$

[89] Nor the President's fleas.

LET'S TURN OUR ATTENTION to an important new idea: *subsets*. The basic idea is that one set, call it A, is a subset of another set, call it B, if every element of A is also an element of B.

8 Definition (Subsets)

We say that A is a subset *of B, written $A \subseteq B$, if and only if for every $a \in A$, we have $a \in B$:*

$$A \subseteq B \iff \forall a \in A\,(a \in B)\,.$$

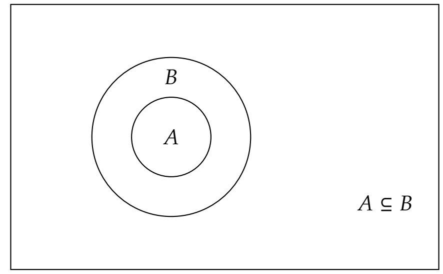


Figure 4: Subset.

Subsets are often introduced with a Venn diagram, as in Figure 4. The rectangle represents the universe of discourse, which is the set that contains everything we care about. Here we see that sets A and B are included in this universe of discourse, and that A is a subset of B. A is a subset of B because every element of A is also an element of B.[90] Observe that every set is a subset of itself, so that $A \subseteq A$ and $B \subseteq B$, too.[91] We can also have sets that are not subsets of each other, as in Figure 5.

[90] Y.H.P. thinks it's useful to think of subsets as "smaller" sets that are *contained* within "larger" sets. More on this in a moment.

[91] Mind binder: can you explain why the empty set is a subset of every set, including itself?

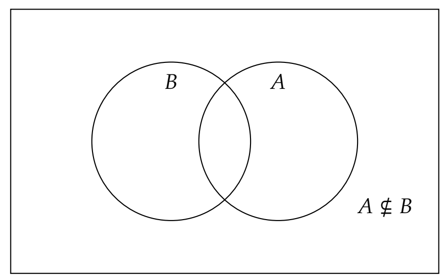


Figure 5: Not a subset.

SUBSETS ARE IMPORTANT because they allow us to organize sets into a hierarchy. To see this, let's introduce a new idea: the *power set of a set.*

9 Definition (Power Set of a Set)

For any set A, the power set *of A, written $\mathsf{P}(A)$, is the set of all subsets of A:*

$$\mathsf{P}(A) = \{B \mid B \subseteq A\}\,.$$

Here are the power sets for the first few counting numbers' worth of elements:

$$\mathsf{P}(\varnothing) = \{\varnothing\}\,,$$

$$\mathsf{P}(\{\heartsuit\}) = \{\varnothing, \{\heartsuit\}\}\,,$$

$$\mathsf{P}(\{\heartsuit, \clubsuit\}) = \left\{ \begin{array}{l} \varnothing, \\ \{\heartsuit\}, \{\clubsuit\}, \\ \{\heartsuit, \clubsuit\} \end{array} \right\},$$

$$\mathsf{P}(\{\heartsuit, \clubsuit, \spadesuit\}) = \left\{ \begin{array}{l} \varnothing, \\ \{\heartsuit\}, \{\clubsuit\}, \{\spadesuit\}, \\ \{\heartsuit, \clubsuit\}, \{\heartsuit, \spadesuit\}, \{\clubsuit, \spadesuit\}, \\ \{\heartsuit, \clubsuit, \spadesuit\} \end{array} \right\},$$

$$\mathsf{P}(\{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\}) = \left\{ \begin{array}{l} \varnothing, \\ \{\heartsuit\}, \{\clubsuit\}, \{\spadesuit\}, \{\diamondsuit\}, \\ \{\heartsuit, \clubsuit\}, \{\heartsuit, \spadesuit\}, \{\heartsuit, \diamondsuit\}, \{\clubsuit, \spadesuit\}, \{\clubsuit, \diamondsuit\}, \{\spadesuit, \diamondsuit\}, \\ \{\heartsuit, \clubsuit, \spadesuit\}, \{\heartsuit, \clubsuit, \diamondsuit\}, \{\heartsuit, \spadesuit, \diamondsuit\}, \{\clubsuit, \spadesuit, \diamondsuit\}, \\ \{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\} \end{array} \right\}.$$

Notice anything about the "size" of the power set? Wait a second, what does "size" even mean?

Again, Y.H.P.?!
Hearts and spades a tired theme
Laughter hides the strain

Mama mia, that's a spicy meatball! Look at all those subsets! Seriously, look at them for a moment before you scroll ahead.[92]

[92] Y.H.P. isn't kidding. Look at them.

Did you?

The power set of a set is a set of sets.[93] It's a set of all the subsets of the original set, and it helps us to see how subsethood creates a hierarchy of sets. Figure 6 shows the power set of {♥, ♣, ♠, ♦}, just as we wrote it out above. The top level contains the original set, and the bottom level contains the empty set. The levels in between contain all the subsets of the original set, and the levels are aligned with the number of elements in the subsets. We draw an arrow from one set to another if the former is a subset of the latter; we don't draw the arrows each and every time.[94]

[93] Yes, the elements of sets can be sets. This is a bit of a mind-bender, but it's also a bit of a mind-bender that the elements of sets can be anything at all. Conditional on the second, the first is not so bad.

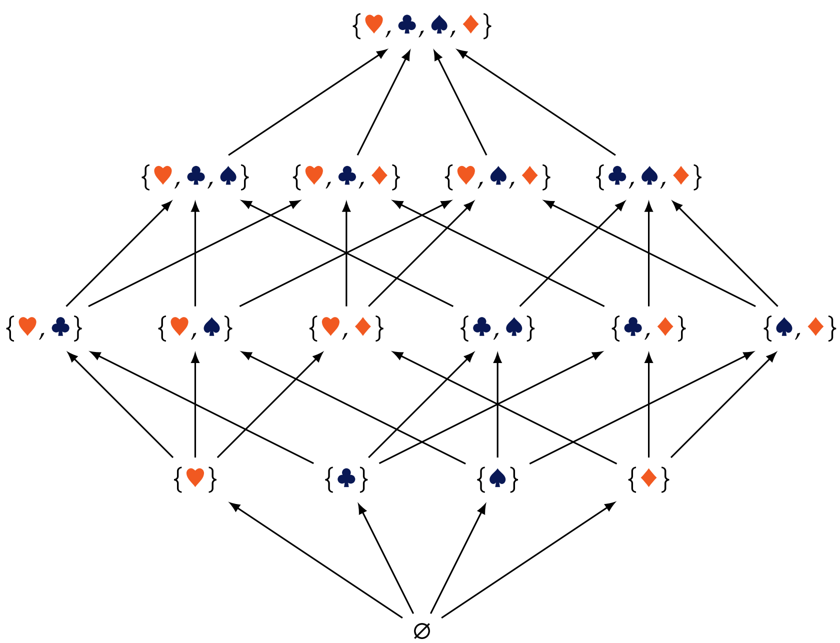


Figure 6: The power set of {♥, ♣, ♠, ♦}.

[94] Do you see which arrows we've omitted? Do you know why that might be? Can you please go check Definition 4 and think about whether it applies to subsethood? And can you please remember that every set is a subset of itself?

This is called a *Hasse diagram.* It's a way of visualizing a hierarchy, in this case the hierarchy of subsets. Take your time figuring out how it works. Hopefully it drives home an important point: it is the relationships between things that matter, not the things themselves. Structure emerges from relationships. The names of things are just labels, and they are, to use a technical term, boring a.f.

If you recreate this diagram for the simpler cases of $P(\varnothing)$, P({♥}), P({♥, ♣}), and P({♥, ♣, ♠}), you'll see that the nature of the hierarchy is the same, and indeed that each smaller hierarchy can be embedded within the larger hierarchy. You'll also see that these are all just fancy ways of drawing high-dimensional cubes on two-dimensional paper. Can you see the "cubes within cubes within cubes" going on in here? Imagine moving some of the nodes around to make it more obvious.

What on earth is the point of all this, at least from the perspective of i.r.? Well, consider the following examples:

1. Suppose ♥, ♣, ♠, and ♦ are four states in the international system. We can model the alliances among them as subsets of the power set of {♥, ♣, ♠, ♦}. For example, the set {♥, ♣} is a subset of the power set, and it represents the alliance between ♥ and ♣.[95]

2. Suppose ♥, ♣, ♠, and ♦ are four pieces of territory. Call each subset a *parcel of land.* Having ∅ would make for a rather small state.

3. Suppose ♥, ♣, ♠, and ♦ are four kinds of actions one state can take against another—say, "diplomacy," "economic sanctions," "military threats," and "military force." Call each subset a *course of action.*

Surely you can think of more examples.

[95] Be flexible here—y.h.p. is just saying that you could, for example, have

♥ = France,
♣ = Germany,
♠ = Italy,
♦ = Spain.

But look how much prettier the suit symbols are! This brings up an interesting question: do you think we should design our models to be pretty? y.h.p. thinks so, but he's a bit of a weirdo. But still: don't you find yourself using things more, and remembering more of their details, if they're beautiful? The people at Apple are banking on it.

Just as with logic, we have operations to let us make new sets from old ones. Indeed, they're even related to the logical operations we saw earlier![96] It's pretty standard just to plow through the next three operations reasonably quickly, and we'll do so now; really, it's only the intersection that we'll use a lot. Notice that all of these operations are defined with respect to some universe of discourse U—that's the $x \in U$ part.

[96] You should be able to convince yourself that subsethood $\subseteq$ is related to logical conditional $\Rightarrow$. But then what maps to $\Leftrightarrow$? We'd need $A \subseteq B$ and $B \subseteq A$ to be true, which is the same as saying $A = B$.

Table 8: Set operations.

Operation	Name	Definition
A^{C}	Complement	$\{x \in U \mid \neg x \in A\}$
$A \cup B$	Union	$\{x \in U \mid x \in A \vee x \in B\}$
$A \cap B$	Intersection	$\{x \in U \mid x \in A \wedge x \in B\}$

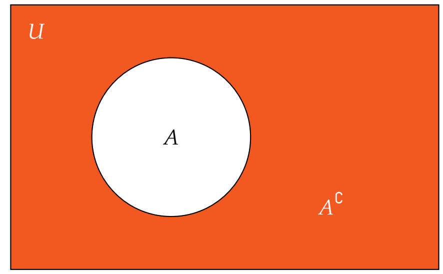


Figure 7: Complement.

The complement of a set A, written A^{C}, is the set of all elements of U that are *not* in A—so as you might have guessed, it's related to the logical negation $\neg$. Formally, we write

$$A^{\mathsf{C}} = \{x \in U \mid \neg x \in A\},$$

which reads "the complement of A is the set of all elements x in the universe of discourse U such that x is not in A." So for example, let U be the set of all states in the international system, $U = \{\text{Afghanistan}, \dots, \text{Zimbabwe}\}$. Let A be the set of all states in Asia, $A = \{\text{Afghanistan}, \dots, \text{Yemen}\}$. Then, A^{C} is the set of all states in the international system that are not in Asia, $A^{\mathsf{C}} = \{\text{Algeria}, \dots, \text{Zimbabwe}\}$.[97] Set complements might seem boring, but they're quite useful. Recalling set notation, we can write

$$A = \{x \in U \mid x \text{ has some property}\},$$

where this might be democracy applied to states, or gender applied to leaders, or competitiveness applied to industries, or anything else you can think of. Then in gathering all the things that bear this property, we're implicitly gathering all the things that don't bear this property, too. To include by a property is to exclude by its negation.[98]

[97] Do you see why we need a universe of discourse? If you were asked what the things not in Asia are, you'd probably say "everything else," but you wouldn't know to what "everything else" refers. As an extreme case, can you say what $\varnothing^{\mathsf{C}}$ is? What about U^{C}?

[98] Right? Like, can you think of a situation where $x \in A$ and $x \in A^{\mathsf{C}}$?

The union of two sets A and B, written $A \cup B$, is the set of all elements that are in either A or B—so as you might have guessed, it's related to the logical disjunction $\vee$. Formally, we write

$$A \cup B = \{x \in U \mid x \in A \vee x \in B\},$$

which reads "the union of A and B is the set of all elements x in the universe of discourse U such that x is in A or x is in B." Continuing the previous example, $A \cup B$ is the set of all states in the international system that are in Eurasia, $A \cup B = \{\text{Afghanistan}, \dots, \text{United Kingdom}, \dots, \text{Yemen}\}$.[99]

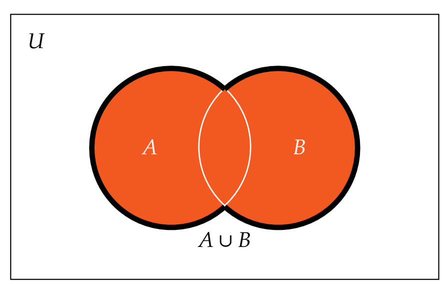


Figure 8: Union.

[99] Riddle y.h.p. this: what are

$$A \cup \varnothing, \quad A \cup U, \quad A \cup A^{\mathsf{C}}, \quad \text{and } A \cup A?$$

Now, for any real number $x \in \mathbb{R}$, what are

$$x + 0, \quad x + \infty, \quad x + (-x), \quad \text{and } x + x?$$

What properties are similar, and what properties are different?

THE INTERSECTION of two sets A and B, written $A \cap B$, is the set of all elements that are in both A and B—so as you might have guessed, it's related to the logical conjunction $\land$. Formally, we write

$$A \cap B = \{x \in U \mid x \in A \land x \in B\},$$

which reads "the intersection of A and B is the set of all elements x in the universe of discourse U such that x is in A and x is in B." Yet again letting U be the set of all states in the international system, again let A be the set of all states in Asia and B the set of all states in Europe. Then $A \cap B$ is the set of all states in *both* Asia and Europe—a set which includes, at least: Russia, Turkey, and Georgia.[100]

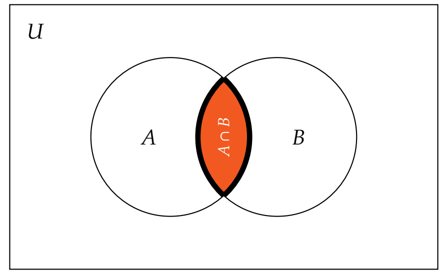


Figure 9: Intersection.

[100] Riddle Y.H.P. that: what are

$A \cap \varnothing$, $A \cap U$, $A \cap A^{\mathsf{C}}$, and $A \cap A$?

Now: which operation for numbers is most similar to intersection for sets? It's kind of strange, but two seemingly-different operations can be quite similar, just with different use cases.

SET INTERSECTION GIVES RISE to an important concept: disjointedness.

---

10 Definition (Disjoint Sets)
We say two sets A and B are disjoint *iff*[101] *their intersection is empty:*

$$A \cap B = \varnothing.$$

Two sets are disjoint if they have no elements in common; in Figure 10, A and B, having an empty intersection, are disjoint.

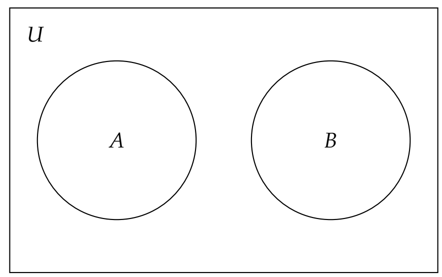


Figure 10: Mutually exclusive.

[101] It's common to see "iff" in mathematical writing. It stands for the phrase "if and only if" we saw earlier. In definitions, Y.H.P. often uses the term "just in case" for the same idea.

In I.R. we refer to pairs of states as *dyads.* And that's what we're identifying here: pairs of distinct states. We have to add the "distinct" part because otherwise we'd have to include sets like {♥, ♥}, which is a set containing a set containing ♥. Now, what sorts of dyadic relationships can we model without appealing to the order of the states in the dyad? After all, sets are unordered.

Let's try to get a feel for this by way of the aforementioned democratic peace thesis. We could assert it in purely logical terms: letting s_1 and s_2 be state variables, we could write

$$\forall s_1 \forall s_2 \left(\mathbf{D}(s_1) \land \mathbf{D}(s_2) \Rightarrow \neg\mathbf{W}(s_1, s_2)\right),$$

where $\mathbf{D}(s_1)$ means "s_1 is a democracy" and $\mathbf{W}(s_1, s_2)$ means "s_1 and s_2 are at war." We can do the same thing with sets. Let U be the set of all states in the international system, and define the following sets:

$$D = \{\{s_1, s_2\} \subseteq U \mid s_1 \neq s_2 \text{ and } s_1 \text{ and } s_2 \text{ are democracies}\},$$
$$W = \{\{s_1, s_2\} \subseteq U \mid s_1 \neq s_2 \text{ and } s_1 \text{ and } s_2 \text{ are at war}\}.$$

Then the democratic peace thesis is that $D \cap W = \varnothing$. But now we can ask some other questions a bit more intuitively, such as: is $D = W^{\mathsf{C}}$—that is, is every democratic dyad peaceful and every peaceful dyad democratic? Or is $D \subseteq W^{\mathsf{C}}$—that is, are all democratic dyads peaceful? It would seem to Y.H.P. that there are peaceful dyads that are not mutually-democratic—wouldn't you agree?[102]

When you think about it, it's kind of amazing that we can take things that we think that we know from a logical perspective, encode them in sets, and then use the logical relationships to figure out how to manipulate the sets. And when you think about it, data is just a bunch of sets, too. So our logical weight-lifting and these definitions give us a way to navigate the world making these data by way of looking at those data![103]

[102] Notice that this brings us back to an interesting point from set-builder notation: if two sets are mutually exclusive, this means that the properties that define them are at odds with each other. We often refer to these properties as mutually exclusive, too, or sometimes as *contradictory.*

[103] How's your head, rookie? You're doing great. Keep pushing. And remember: this isn't a book, and you don't have to read it like one. You can skip around, and you can come back to things later. Y.H.P. won't be offended.

Time for a left turn: have you ever seen a *scatterplot?* It's a way of visualizing data. Consider Figure 11, which shows the relationship be-

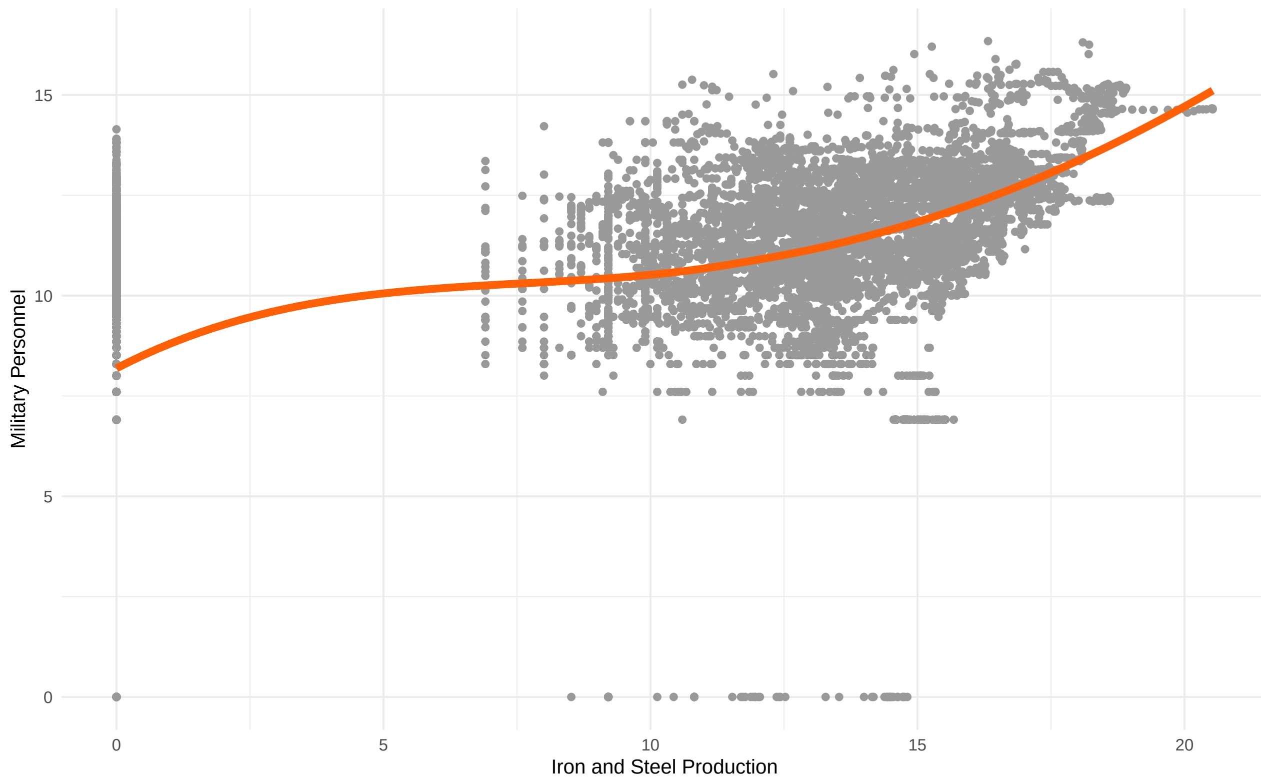


Figure 11: The relationship between iron and steel production and military personnel. Data from the Correlates of War project.

tween iron and steel production and military personnel for all states in the international system.[104]

But what is going on here, exactly? What are these dots, and what do they represent? And once we've figured that out, what is the appropriate container to store such dots in? Clearly the exercise is similar to what you've seen in grade school: putting points in a plane. For example, we have the observation for the United States in 1816:

$$(11.289, 9.740)$$

If we wanted to find the dot for the United States in 1816, we'd move 11.289 units along the x axis and 9.740 units along the y axis. The same goes for all of the observations in the data set. So, we're saying that we can describe a point in a plane by two numbers, which we call the x and y coordinates. Notice that the order of the coordinates matters: if we switched them, we'd be in a different place:[105]

$$(9.740, 11.289) \neq (11.289, 9.740)\,.$$

But we don't yet have a container that lets order matter!

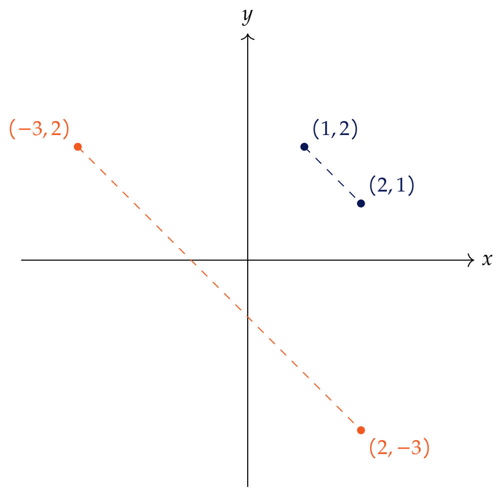
[104] For the record, these values have been logged to make it easier to see the relationship. You can see some sample data below: military personnel is in thousands, and iron and steel production is in thousands of tons.

stateabb	year	irst	milper
USA	1816	11.289	9.740
USA	1817	11.289	9.615
–	–	–	–
USA	2015	18.177	14.172
USA	2016	18.175	14.139
NIG	1960	0.000	8.517
NIG	1961	0.000	8.699
–	–	–	–
NIG	2015	11.513	11.389
NIG	2016	11.513	11.389
CHN	1860	9.210	13.815
CHN	1861	9.210	13.815
–	–	–	–
CHN	2015	20.506	14.660
CHN	2016	20.506	14.660

[105] The latitude and longitude for our classroom is $(40.10478°N, -88.23188°E)$. If we flipped those numbers to $(-88.23188°N, 40.10478°E)$, we'd be deep, deep in Antarctica. Suddenly Urbana sounds nice for once.

We can fix this by using *ordered pairs*. An ordered pair is—you guessed it!—a pair of things, where the order matters.[106] We write an ordered pair as (x, y), where x and y are the things in the pair. If given two ordered pairs (x_1, y_1) and (x_2, y_2), we say they're equal iff $x_1 = x_2$ and $y_1 = y_2$. So, for example, $(1,2) \neq (2,1)$, but $(1,2) = (1,2)$.

[106] For this, you're paying.



So this brings us to a very important construction: the *Cartesian product* of two sets. Definitions first:

---

11 Definition (Cartesian Product)

Let A and B be sets. Then the Cartesian product *of A and B, written $A \times B$, is the set of all ordered pairs (a, b) such that $a \in A$ and $b \in B$:*

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

In particular, we have the Cartesian square *of A, written A^2, which is the Cartesian product of A with itself:*

$$\begin{aligned} A^2 &= A \times A, \\ &= \{(a_1, a_2) \mid a_1 \in A \text{ and } a_2 \in A\}. \end{aligned}$$

For example, suppose $S = \{♥, ♣, ♠, ♦\}$ is the set of suits in a deck of cards, and $R = \{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\}$ is the set of ranks in a deck of cards. Then we can use the two dimensions afforded by these sets to describe every card in the deck in a tidy way:

	2	3	4	5	6	7	8	9	10	J	Q	K	A
♥	2♥	3♥	4♥	5♥	6♥	7♥	8♥	9♥	10♥	J♥	Q♥	K♥	A♥
♣	2♣	3♣	4♣	5♣	6♣	7♣	8♣	9♣	10♣	J♣	Q♣	K♣	A♣
♠	2♠	3♠	4♠	5♠	6♠	7♠	8♠	9♠	10♠	J♠	Q♠	K♠	A♠
♦	2♦	3♦	4♦	5♦	6♦	7♦	8♦	9♦	10♦	J♦	Q♦	K♦	A♦

You might be wondering about the relevance of this to I.R.[107] Consider the implications of order in different relationships:

Do notice that we're taking advantage of dimensionality here: we're using S to set the terms for rows and R to set the terms for columns. What sets are "setting the terms" in Figure 11?

[107] How dare you ask reasonable questions?!

1. Trade: Is it just that State s_1 trades with State s_2, or does s_1 specifically export to s_2?
2. War: Is it merely that s_1 is at war with s_2, or did s_1 initiate the attack on s_2?
3. Alliances: Are States s_1 and s_2 allies, or is s_1 the patron of s_2?
4. ICJ Disputes: Are States s_1 and s_2 simply in a dispute, or is s_1 specifically the complainant against s_2 at the ICJ?

Notice that some of these relationships seem to imply an asymmetry, too: for example, it's hard to imagine s_1 being both the patron and the protege of s_2. However, it's easy to imagine s_1 both importing from and exporting to s_2.

Surely you can think of more examples, but for each of them we can see that ordered information sometimes lurks beneath seemingly-unordered data.[108] In other words, we can think about whether a given relationship is symmetric (like being related) or asymmetric (like being ancestors).

[108] And there's more information in the directed data: if you give me both the imports and the exports, I can tell you the trade balance by subtracting the former from the latter. But if you give me the trade balance, I can't back out the imports and exports, because many import-export pairs can give rise to the same trade balance.

So just to bring this back to the plane, we can think of the Cartesian product as a way of describing things. In particular, we have

$$\begin{aligned} \mathbb{R}^2 &= \mathbb{R} \times \mathbb{R}, \\ &= \{(x_1, x_2) \mid x_1 \in \mathbb{R} \text{ and } x_2 \in \mathbb{R}\}\,. \end{aligned}$$

As you can see at right, this provides a way of describing points in the plane: this is how we can put iron and steel production and military personnel in a plane, dot by dot by blessed little dot.[109]

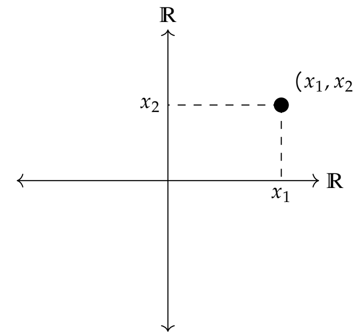


Figure 12: The Cartesian plane.

[109] Bless your heart, little dot!

So what's stopping us from doing this for three dimensions? For example, in Figure 11, we don't take advantage of all the information at hand: for example, we know the year of each observation, but we don't use it. For example, observation for the United States in 1816 could be

$$(11.289, 9.740, 1816)\,.$$

But that's three numbers! Well how do you feel about this?

$$\begin{aligned} \mathbb{R}^3 &= \mathbb{R} \times \mathbb{R} \times \mathbb{R}, \\ &= \{(x_1, x_2, x_3) \mid x_1 \in \mathbb{R} \text{ and } x_2 \in \mathbb{R} \text{ and } x_3 \in \mathbb{R}\}\,. \end{aligned}$$

Yes, we could be a little more specific about what a year is—after all, negative time isn't allowed. But it's just easier to say "just let that be another real number" so that Y.H.P. can teach you the notation $\mathbb{R}^3$.

This provides a home for all the triples of real numbers, and we can use it to describe points in three-dimensional space. Sadly, Y.H.P. can't draw a picture of this for you, but you can imagine it.

Naturally, we can do this for any number of dimensions, so that n-dimensional space is $\mathbb{R}^n$:

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_1 \in \mathbb{R} \text{ and } x_2 \in \mathbb{R} \text{ and } \cdots \text{ and } x_n \in \mathbb{R}\}\,.$$

If we think about quantitative data as real numbers, then we can think about data as points in n-dimensional space. $\mathbb{R}^n$ thus allows us to think about data in a very general way.[110]

[110] While it's true that one can learn how to analyze data without knowing any of this, it's also true that one can learn how to drive a car without knowing how to change the oil. But now and again, it's nice to know how to change the oil.

But remember: Cartesian products can combine any sets, not just numeric ones. For example, suppose had a set of states $\{s_1, \ldots, s_n\}$, and suppose each could vote Yea, Nay, or Present for a resolution on the floor of the United Nations General Assembly. Then we could describe the votes of all states on the resolution as

$$(v_1, \ldots, v_n) \in \{Y, N, P\}^n\,,$$

where each v_i is the vote entered by state s_i. For example, we have the three unanimous voting profiles

$$(Y, \ldots, Y), \quad (N, \ldots, N), \quad \text{and } (P, \ldots, P),$$

each of which is an element of $\{Y, N, P\}^n$.[111] We therefore have a way of describing all possible voting profiles for n states on a resolution.

Now ask yourself: for the purposes of voting, does the order of the votes matter? Or could you just calculate a tally of Yea, Nay, and Present votes? Put differently: does it matter which state enters which vote, or are all Yeas, Nays, and Presents the same? Put a star next to this one: we'll consider a similar question in a few weeks.

[111] But we also have the profile

$$(Y, N, Y, \ldots, Y) \in \{Y, N, P\}^n\,,$$

which is not unanimous. Each and every combination of votes is an element of $\{Y, N, P\}^n$.

Relations

We've seen that sets can be used to describe all sorts of things. But what if we want to describe the relationships between things? For example, suppose we wanted to say something like

State s_1 is at least as powerful as State s_2,

where s_1 and s_2 are variables ranging over states in the international system.[112] Now, we could define this statement in purely logical terms, introducing a new predicate **P** for "is at least as powerful as," then writing

$\mathbf{P}(s_1, s_2) \leftrightsquigarrow$ "State s_1 is at least as powerful as State s_2."

But we could also define this statement in terms of sets; this would be great, because then we'd have more machinery to work with.

It turns out that we can encode relations in terms of sets.

[112] Y.H.P. will now be casual in folding in new variables; you'll get an ear for it.

12 Definition (Binary Relations)

Let A and B be sets. Then a binary relation *R from A to B is a subset of the Cartesian product* $A \times B$*:*

$R \subseteq A \times B.$

Notice that it's not just practical considerations driving the pursuit of set-theoretic definitions of basic things; it's also the fact that set theory forms the basis of mathematics, so it is desirable to define things in terms of sets.

For example, suppose we had the set $S = \{♥, ♣, ♠, ♦\}$. We can envision its Cartesian square as a nice little table like the one at right. Now, suppose we wanted to describe the relationship "is the same color as." Well, Definition 12 tells us that we need to find a subset of $\{♥, ♣, ♠, ♦\}^2$ that captures the essence of this relationship. Well, which pairs of suits are the same color? It's just the pairs of suits that are both red or both black. So we can define the relationship "is the same color as" as the subset of $\{♥, ♣, ♠, ♦\}^2$ that contains all the pairs of suits that are the same color. This subset is shown in Table 10. Observe that this subset is a relation from $\{♥, ♣, ♠, ♦\}$ to $\{♥, ♣, ♠, ♦\}$, because it's a subset of $\{♥, ♣, ♠, ♦\}^2$.[113] And, the complement of this relation is the relation "is a different color than," which is shown in Table 11.

	♥	♣	♠	♦
♥	(♥,♥)	(♥,♣)	(♥,♠)	(♥,♦)
♣	(♣,♥)	(♣,♣)	(♣,♠)	(♣,♦)
♠	(♠,♥)	(♠,♣)	(♠,♠)	(♠,♦)
♦	(♦,♥)	(♦,♣)	(♦,♠)	(♦,♦)

Table 9: $\{♥, ♣, ♠, ♦\}^2$.

[113] We say a relation is *symmetric* just in case any time (a, b) is in the relation, (b, a) is in the relation, too. Do you think "is the same color as" is symmetric?

	♥	♣	♠	♦
♥	(♥,♥)			(♥,♦)
♣		(♣,♣)	(♣,♠)	
♠		(♠,♣)	(♠,♠)	
♦	(♦,♥)			(♦,♦)

Table 10: "Is the same color as."

	♥	♣	♠	♦
♥		(♥,♣)	(♥,♠)	
♣	(♣,♥)			(♣,♦)
♠	(♠,♥)			(♠,♦)
♦		(♦,♣)	(♦,♠)	

Table 11: "Is a different color than."

Though we define relations in terms of sets, there remains considerable mathematical and philosophical controversy about the nature of relations. The most strident of the axiomatic type of thinkers will tell you that relations are sets of ordered pairs, full stop—after all, sets are all there is, and progress is simply of matter of building new architecture on top of the foundation of sets. To others, relations are the encoding of the essence of relationships between things, and sets are just a convenient way of describing them. Put differently, relations are rules we come up with to describe the relationships between things, and sets are just a convenient way of describing those rules. Y.H.P. is in the latter camp, but it's nice to know that there's a debate here.

IN THE PREVIOUS EXAMPLE, our relation linked elements of the same set. But there's no reason why we can't link elements of different sets. For example, consider the relation

Individual i is the leader of State s.

Well jeez, that's different.[114] Let's define two sets:

[114] For this, you're paying.

$$I = \{i_1, \dots, i_{n_I}\},$$
$$S = \{s_1, \dots, s_{n_S}\},$$

where I is the set of all individuals on Earth right now and S is the set of all states in the international system. These are both finite sets; there are $n_I \in \mathbb{N}$ individuals and $n_S \in \mathbb{N}$ states.[115] It would be very annoying to try to write this out as a table, and yet we can easily write out the relevant Cartesian product:

[115] So what are those numbers? Well, there are about 8.1 billion people on Earth right now, so $n_I \approx 8.1 \times 10^9$. And there are about 200 states in the international system, so $n_S \approx 200$.

Notice that we're "dot dot dot"ting twice here: once for the individuals and once for the states. So we've got $(i_1, s_1), \dots, (i_{n_I}, s_1)$, which is every individual against the first state; that's the individual-level dot-dot-dot. And then we repeat that exercise for every state, so we've got $(i_1, s_1), \dots, (i_{n_I}, s_1), \dots, (i_1, s_{n_S}), \dots, (i_{n_I}, s_{n_S})$, which is the state-level dot-dot-dot.

$$\begin{aligned} I \times S &= \{(i, s) \mid i \in I \text{ and } s \in S\}, \\ &= \{(i_1, s_1), \dots, (i_1, s_{n_S}), \dots, (i_{n_I}, s_1), \dots, (i_{n_I}, s_{n_S})\}. \end{aligned}$$

Now, we can define the relation "is the leader of" as the subset of $I \times S$ that contains all the pairs of individuals and states that are linked by the relationship "is the leader of." What are some elements of this set? Well, at the time of this writing, we've got

$$\begin{aligned} (\text{Donald Trump}, \text{United States}) &\in \{\text{leaders}\} \subseteq I \times S, \\ (\text{Xi Jinping}, \text{China}) &\in \{\text{leaders}\} \subseteq I \times S, \\ (\text{Bola Tinubu}, \text{Nigeria}) &\in \{\text{leaders}\} \subseteq I \times S, \\ (\text{Giorgia Meloni}, \text{Italy}) &\in \{\text{leaders}\} \subseteq I \times S, \end{aligned}$$

And we *definitely* don't have

$$(\text{United States}, \text{Donald Trump}) \in \{\text{leaders}\} \subseteq I \times S,$$

because the order matters. If you wanted that, you'd need the relation "is led by." Now, if you were thinking about these as tables like we did with the suits, you might have a quick way of turning the table for "is the leader of" into the table for "is led by." Can you see it? Get creative!

and so on. But notice that we don't have

$$\begin{aligned} (\text{Donald Trump}, \text{China}) &\in \{\text{leaders}\} \subseteq I \times S, \text{ nor} \\ (\text{Y.H.P.}, \text{United States}) &\in \{\text{leaders}\} \subseteq I \times S, \end{aligned}$$

because Donald Trump is not the leader of China, and Y.H.P. is not the leader of the United States.[116] So as you can see, we need not just relate things in the same set; we can relate things in different sets, too.

[116] Thank the Almighty.

Of course, one can argue about what is meant by "is the leader of." Y.H.P. was using a "head of government" criterion, but one could say that Individual i is a leader of State s if i is either the head of government or the head of state. Using that criterion, we'd have both

$$\begin{aligned} (\text{Giorgia Meloni}, \text{Italy}) &\in \{\text{leaders}\} \subseteq I \times S, \\ (\text{Sergio Mattarella}, \text{Italy}) &\in \{\text{leaders}\} \subseteq I \times S, \end{aligned}$$

as the former is the head of government and the latter is the head of state. One object (here Italy) can be connected to many other objects (here Giorgia Meloni and Sergio Mattarella), and many objects (here Giorgia Meloni and Sergio Mattarella) can be connected to one object (here Italy).[117]

[117] Does this remind you of anything? Say, vertical line tests of yesteryear?

SINCE CARTESIAN PRODUCTS can be iterated to allow for many sets, it's natural to wonder whether relations can be similarly extended. Well, consider the following relation:

Knowing facts is neat. But knowing how to use facts is neater. And knowing how to think about how to use facts is neatest. So while Y.H.P. is impressed that you know that the Sultan of Brunei's name is Hassanal Bolkiah and that he's been in power since 1967, he'd be even more impressed if you could use that information—probably in tandem with other information—to answer a question about the international system.

The leader of State s in Year y is Individual i.

Now we've got three things to relate: states, years, and individuals. So we can define three sets:

$$
\begin{aligned}
I &= \{i_1, \ldots, i_{n_I}\}, \\
S &= \{s_1, \ldots, s_{n_S}\}, \\
Y &= \{y_1, \ldots, y_{n_Y}\},
\end{aligned}
$$

where I is the set of all individuals on Earth right now, S is the set of all states in the international system, and Y is the set of all years in the Common Era.[118] Now consider the set $S \times Y \times I$, which is the Cartesian product of S, Y, and I. Now it's impossible to use dimensionality to simplify this, because we've got three sets and two-dimensional paper. But we can still write out the elements of $S \times Y \times I$; an arbitrary one would look like (s, y, i). Now if we define the historical-leaders relation as the subset of $S \times Y \times I$ that contains all the triples of states, years, and individuals that are linked by the relationship "is the leader of," then we can write out some elements of this relation:

[118] The $S \times Y$ part—which is to say, the states and the years—play an important part in data-driven I.R. research. One sometimes sees the altogether disgusting term "spatiotemporal domain" to describe this set. All that is is: the set of state-year combinations under consideration. For many studies, the years range from 1816 to the present, 1816 being the start of a new European state system after the Napoleonic Wars.

$$
\begin{aligned}
(\text{United States}, 1789, \text{George Washington}) &\in \{\text{historical-leaders}\} \subseteq S \times Y \times I, \\
&\vdots \\
(\text{United States}, 2026, \text{Donald Trump}) &\in \{\text{historical-leaders}\} \subseteq S \times Y \times I, \\
&\vdots \\
(\text{India}, 1947, \text{Jawaharlal Nehru}) &\in \{\text{historical-leaders}\} \subseteq S \times Y \times I, \\
&\vdots \\
(\text{India}, 2026, \text{Narendra Modi}) &\in \{\text{historical-leaders}\} \subseteq S \times Y \times I,
\end{aligned}
$$

and so on. So, if we want to be a little more general about things, we can define an n-ary relation like so:

---

13 Definition (Finite Relations in General)

*Let $A_1, \ldots, A_n$ be sets. Then an n-*ary relation *R from $A_1, \ldots, A_n$ is a subset of the Cartesian product $A_1 \times \cdots \times A_n$:*

$$R \subseteq A_1 \times \cdots \times A_n.$$

Indeed, one can take the Cartesian product of an infinite number of sets. But we'll never need to do that, and it introduces some serious complications. In case you're ever bored, look up the *axiom of choice* and the *well-ordering theorem*.

In other words, we can come up with very rich descriptions of relationships between things, and we can use sets to describe those relationships. And since we had agreed earlier that relationships between things are more interesting than things, we can use sets to describe the most interesting things of all, even all they do is gather things up: they are the tool, and we are the craftspeople.[119]

[119] How is this both empowering and terrifying at the same time?!

WHILE WE'RE HERE, let's record a few properties of binary relations. These will prove important in the coming weeks, and they'll help you develop comfort with the machinery.

---

14 Definition (Properties of Binary Relations)

Let A be a set, and let R be a binary relation from A to A. We say that R is:

1. reflexive *just in case for all $a \in A$, $(a, a) \in R$;*
2. irreflexive *just in case for all $a \in A$, $(a, a) \notin R$;*
3. symmetric *just in case for all $a, b \in A$, if $(a, b) \in R$, then $(b, a) \in R$;*
4. antisymmetric *just in case for all $a, b \in A$, if $(a, b) \in R$ and $(b, a) \in R$, then $a = b$;*
5. asymmetric *just in case for all $a, b \in A$, if $(a, b) \in R$, then $(b, a) \notin R$;*
6. complete *just in case for all $a, b \in A$, either $(a, b) \in R$ or $(b, a) \in R$;*
7. transitive *just in case for all $a, b, c \in A$, if $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$.*

Don't. Freak. Out. Y.H.P. is just trying to get you to think about the properties of relations. The tedium of a given definition is *not* a reflection of its importance. Don't lose sight of the forest for the trees.

In one of Y.H.P.'s favorite math books, this laundry list of properties is referred to as a "bestiary of relations." Unleash the beast in your mind, dear reader. Unleash it good.

Let's think some of these through.

1. *Reflexive* means that every element of A is related to itself. Let A be the set of all individuals on Earth right now. Then the relation "is the same age as" is reflexive, because every individual is the same age as themselves. Conversely, the relation "is older than" is not reflexive, because no individual is older than themselves—indeed, we can say that "is older than" is irreflexive.[120]

2. *Symmetric* means that if a is related to b, then b is related to a. Again let A be the set of all individuals on Earth right now. Then the relation "is related to by blood at at least the second degree" is symmetric, because if a is related to b by blood at at least the second degree, then b is related to a by blood at at least the second degree.[121] Conversely, the relation "is the parent of" is not symmetric, because if a is the parent of b, then b is not the parent of a—indeed, we can say that "is the parent of" is asymmetric.

3. *Antisymmetric* means that if a is related to b and b is related to a, then a and b are the same thing. This is a bit trickier to do with individuals, so let's consider numbers: let $A = \mathbb{N}$, so that we're dealing with counting numbers, and consider the relation $\geqslant$. Now, it's possible for two numbers a_1 and a_2 to have the property that $a_1 \geqslant a_2$ and $a_2 \geqslant a_1$. But if that's the case, then $a_1 = a_2$.[122] So $\geqslant$ is antisymmetric: the only way for "symmetry" to hold is if the two things are the same thing. Can you think of a substantive example of an antisymmetric relation? How about a relation that is not antisymmetric?

[120] So here's a question: what about the relation "is genetically identical to"? Is that reflexive? Is it irreflexive? In case you're confused, throw on H.G.T.V. and wait for an episode of *Property Brothers*.

[121] "At least the second degree" means that a and b are either siblings, parent and child, or grandparent and grandchild.

[122] Conversely, $>$ is asymmetric, because if $a_1 > a_2$, then $a_2 \not> a_1$. And $=$ is symmetric, because if $a_1 = a_2$, then $a_2 = a_1$. Antisymmetry is an important part of our notion of order, but it's still a bit tricky to wrap one's head around.

4. *Complete* means that for any two elements of A, the relationship holds in at least one direction.[123] Let A be the set of all individuals on Earth right now. Then the relation "is taller than" is complete, because for any two individuals, one is taller than the other. If one is strictly taller than the other, then the relation holds in one direction; if they're the same height, then the relation holds in both directions. Conversely, the relation "is the parent of" is not complete, because one can find a pair of individuals—say, you and Y.H.P.—such that neither is the parent of the other. One way of thinking about complete relations is that they make all elements of A *comparable* to each other in some sense.[124]

5. We've already seen *transitivity* in action. Let A be the set of all individuals on Earth right now. The relation "is the ancestor of" is transitive, because if a is the ancestor of b and b is the ancestor of c, then a is the ancestor of c.[125] However, "is the parent of" is not transitive, because if a is the parent of b and b is the parent of c, then a is not the parent of c—a is the grandparent of c. Now let's take this sidenote at right very seriously: let S be the set of all states in the international system, and suppose you think s_1 is more powerful than s_2 and s_2 is more powerful than s_3. Can you conclude that s_1 is more powerful than s_3? Well, how do you define "more powerful than"? If it's "would win a war against," then perhaps not: maybe there's something about the s_1-s_2 war that makes s_1 more powerful than s_2, but that something doesn't apply to the s_1-s_3 war.[126] So if you want to say that "is more powerful than" is transitive, you need to be very careful about how you define "is more powerful than." Somebody that uses the relation as if it's transitive is making some very strong assumptions about the nature of power. That might be perfectly fine, but it's important to be aware of it.[127]

As a final quick definition, let's define equivalence relations. Notice that the word "equivalence" is in there, so we're talking about relations that are somehow "equal" to each other. What is it about = that makes it so special? What properties does it have that we think might apply to other similar relations?

---

15 Definition (Equivalence Relations)

Let A be a set and let R be a binary relation from A to A. We say R is an equivalence relation *just in case R is reflexive, symmetric, and transitive.*

Let's confirm that = is indeed an equivalence relation on $\mathbb{R}$: it is reflexive, as any number $x \in \mathbb{R}$ satisfies $x = x$; it is symmetric, as if $x = y$, then $y = x$; and it is transitive, as if $x = y$ and $y = z$, then $x = z$. So = is an equivalence relation on $\mathbb{R}$.[128] Now, would you say that "is the same age as" is an equivalence relation on the set of all individuals on Earth right now? How about "is as powerful as" on the set of all states?

[123] What can you say about the relationship between completeness and reflexivity? *Hint:* you can say something. *Hint hint:* you can say that one implies the other. *Hint hint hint:* think about a toy example with a set of two elements.

[124] Wow, comparability is such a cool thing to think about. It's a very important part of the structure of the international system. There are some ways in which states are comparable, and some ways in which they are not. And there are some ways where it's not clear, and these are often the most fun to think about. For example, would you say the relation "is more powerful than" is complete? If so, then this means that all states can be identified by some underlying scale of power, and that these scales are comparable. But if not, then this means that there are some states that are not comparable in terms of power. Does it matter what kind of power we're talking about? Can you compare Vatican City's power to North Korea's power?

[125] This is kind of fun to think about: every time you get a transitive relation like "is the ancestor of," you get its opposite, too, and you know it's transitive. Thus, if "is the ancestor of" is transitive, then so too is "is the descendant of."

[126] Or consider basketballsing: if the Illini beat the Hoosiers and the Hoosiers beat the Boilermakers, does it follow that the Illini beat the Boilermakers? That seems to depend on the matchups, doesn't it? Purdue always has high-quality big men. And China always has labor-intensive armies!

[127] Assmptions are cool—indeed, they're necessary. But assumptions you're not aware of can be dangerous, particularly if they influence the concepts you use to think about the world. So be aware of your assumptions: this is one reason to adopt a structured approach to encoding, rather than just winging it at a dinner party.

[128] But what about $\geq$? Well, it's reflexive, because any number $x \in \mathbb{R}$ satisfies $x \geq x$. And it's transitive, because if $x \geq y$ and $y \geq z$, then $x \geq z$. But it's not symmetric, because it could well be that $x \geq y$ but not $y \geq x$—indeed, this is the case for $x = 2$ and $y = 1$. So $\geq$ is not an equivalence relation on $\mathbb{R}$.

Functions

We've seen that sets can be used to describe all sorts of things, and we've seen that relations can be used to describe the relationships between things. But there are certain kinds of relationships that are especially important; let's call these *functional relationships*.

---

16 Definition (Functions)

Let A and B be sets. Then a function *f from A to B is a binary relation from A to B such that for all $a \in A$, there exists a unique $b \in B$ such that $(a, b) \in f$. In such cases, we write $f : A \to B$, calling A f's* domain *and B f's* codomain.[129]

---

[129] You might have heard B called f's *range*, but that's not quite right. It's safer to call it the codomain, for reasons Y.H.P. can explain to you if you're interested.

Functions are a reliable sort of relationship: for every element of A, there's exactly one element of B that it's related to. In particular:

1. every element of a is sent to *something* in B, meaning that the relationship is well-defined enough that we can talk about it; and
2. every element of a is sent to *exactly one* thing in B, meaning that the relationship is well-defined enough that we can talk about it *uniquely*.

You'll find a sample function in Figure 13. Just to make sure we're clear here, consider non-examples as depicted in Figures 14 and 15.

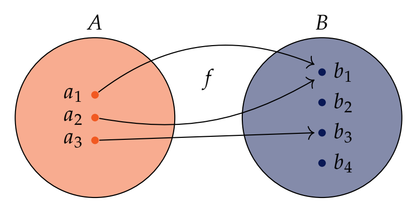


Figure 13: A function $f : A \to B$.

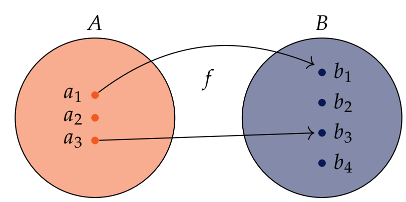


Figure 14: Not a function, because a_2 is not sent to anything.

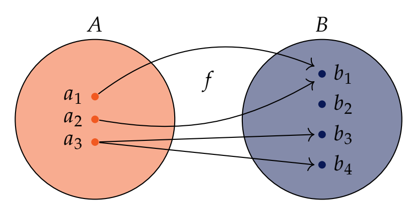


Figure 15: Not a function, because a_3 is sent to more than one thing.

Functions are important because they allow us to describe how things change. Let S be the set of all states in the international system, and let Y be the set of all years in the Common Era. Then we can define a function $f : S \times Y \to \mathbb{N}$ that maps a state-year pair to the number of people in that state in that year. Now, this is a very simple function, but it's still a

$s \in S$	$y \in Y$	$(s, y) \in S \times Y$		$f(s, y) \in \mathbb{N}$
USA	2020	(USA, 2020)	⟶	328,239,523
CHN	2020	(CHN, 2020)	⟶	1,439,323,776
IND	2020	(IND, 2020)	⟶	1,380,004,385
RUS	2020	(RUS, 2020)	⟶	145,934,462
BRA	2020	(BRA, 2020)	⟶	212,559,417
USA	2010	(USA, 2010)	⟶	308,745,538
CHN	2010	(CHN, 2010)	⟶	1,341,335,000
IND	2010	(IND, 2010)	⟶	1,210,193,422
RUS	2010	(RUS, 2010)	⟶	142,958,164
BRA	2010	(BRA, 2010)	⟶	190,732,694
⋮	⋮	⋮		⋮

function. For every state-year pair, there's exactly one number of people in that state in that year—not two, not zero, not a million, but exactly one. And that's what makes it a function.[130]

[130] Notice the usefulness of the Cartesian product here: it's the thing that lets the function depend on multiple inputs, giving rise to what one might call *multivariate functions*. In game theory, all utility functions are defined on Cartesian products, so put a star next to this paragraph.

That said, some seemingly-straightforward relationships are not functions. For example, let S be the set of all states in the international system, and let C be the set of all cities in the world. Define the relation R as

The capital city of State s is City c.

Now, this is a perfectly reasonable relationship to consider—you probably considered it several times during grade school—but it's not a function. Why not? Well, it turns out that some states have multiple capital cities; see Table 12. So if we want to define a function that maps states to capital cities, we need to be careful about how we define the domain and codomain, or else we wind up with something that looks like Figure 15.[131]

State	Capital City	Function
South Africa	Pretoria	Executive
South Africa	Cape Town	Legislative
South Africa	Bloemfontein	Judicial
Netherlands	Amsterdam	Constitutional
Netherlands	Den Haag	Seat of Government
Bolivia	Sucre	Constitutional
Bolivia	La Paz	Seat of Government

Table 12: Selected states with multiple capital cities.

[131] Try and draw the problem in your own blobs-dots-and-arrows diagram. It's a good exercise. Meanwhile, some states only have *de facto* capital cities with no formal recognition—say, Switzerland, which has no capital city but rather a *de facto* capital in Bern. Thus, if we define the rule by which we assign each state its formal capital city, we'd bump into a Figure 14 problem: a dot (Switzerland) with no arrow. Wha happen?!

Like relations, functions can have certain properties. There are two that are espeially relevant for our current purposes.

17 Definition (One-to-One and Onto Functions)

Let $f : A \to B$ be a function. Then f is:

1. one-to-one *(or* injective*) just in case for all $a_1, a_2 \in A$, if $f(a_1) = f(a_2)$, then $a_1 = a_2$; and*
2. onto *(or* surjective*) just in case for all $b \in B$, there exists an $a \in A$ such that $f(a) = b$.*

In case a function is both one-to-one and onto, we say it is bijective*.*

Let's think these definitions through.[132]

1. *One-to-one* means that each element of A gets sent to its own special element of B. The only time two things get sent to the same place is if they're the same thing—that's the formal definition above. A function that's one-to-one is fully faithful to the elements of its domain: it doesn't conflate them, it doesn't confuse them, it doesn't mix them up. It doesn't pixelate them, it doesn't blur them, it doesn't distort them. Every unique thing remains unique after the function is applied, kind of like embedding a photograph in a PDF.
2. *Onto* means that every element of B gets sent to by *something* in A. The function is fully faithful to the elements of its codomain: it doesn't leave any of them out, it doesn't forget any of them, it doesn't ignore any of them. It doesn't crop them, it doesn't cut them, it doesn't erase them. We said that we needed B to house the range of the function, and we've done this in a way that didn't leave any extra space in B.

Notice that one-to-one functions need not be onto, and onto functions need not be one-to-one. However, a function can indeed bear both properties, as depicted in Figure 18. This is a function that neither pixelates nor crops: it's fully faithful to both the domain and the codomain. Yee haw, that's a bijective function!

[132] The function depicted in Figure 13 is not one-to-one, because a_2 and a_3 are both sent to b_3. Neither is it onto, because b_4 is not sent to by any element of A. So consider the drawings below "modifications" of Figure 13 that make it one-to-one or onto.

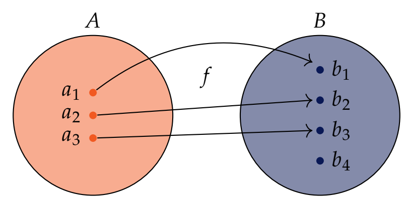


Figure 16: A one-to-one (but not onto) function $f : A \to B$.

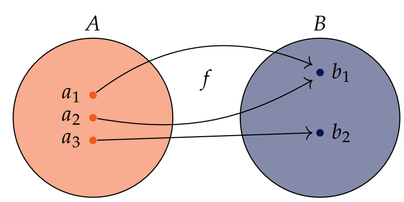


Figure 17: An onto (but not one-to-one) function $f : A \to B$.

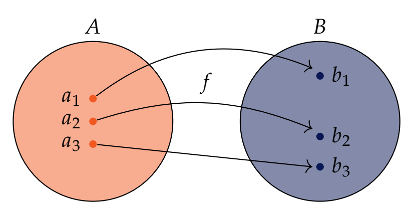


Figure 18: A one-to-one and onto function $f : A \to B$.

You may have noticed that one-to-one and onto seem to be related to the number of elements in the domain and codomain. But, we haven't yet defined what it means for a set to have a certain number of elements. Let's remedy that—to the Definitions Cave!

You might be tempted to think that the cardinality is a function $|\cdot| : \{\text{sets}\} \to \mathbb{N} \cup \{0\}$ that maps a set to the number of elements it contains. But there's a problem with this: the domain of the function is the set of all sets, and that's not a set. It's a *proper class*, which is a collection of things that's too big to be a set. So we can't define a function whose domain is the set of all sets.

---

18 Definition (Cardinality of Sets)

The cardinality *of a set A, denoted $|A|$, is the number of elements in A.*

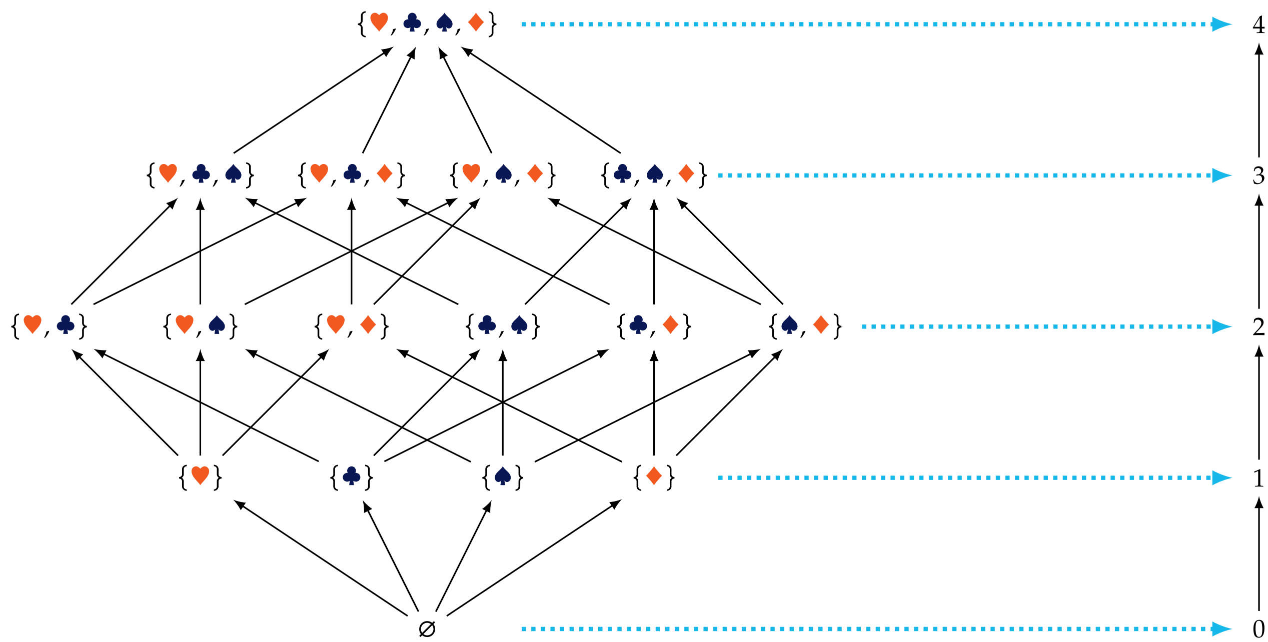


Figure 19: Set cardinality for $\mathcal{P}(\{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\})$.

Figure 19 shows the cardinality for the elements the power set of $\{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\}$. At the bottom we see the empty set, which (having no elements) is sent to 0. Next we see the four one-suit sets,[133] which are sent to 1. And so on, until we hit the full set, which is sent to 4.[134]

There's a bit of a curious thing going on here: notice that there are four sets with cardinality 1, namely $\{\heartsuit\}$, $\{\clubsuit\}$, $\{\spadesuit\}$, and $\{\diamondsuit\}$. These sets are all pairwise disjoint—they're all different sets, right? They don't have any elements in common, and yet there's something that they *do* have in common: they all have exactly one element. So they have something in common, but they're not the same. And yet, there's the same *in a sense*, because they all have the same cardinality. Let's write $A \Bumpeq B$ in case sets $|A| = |B|$. Then we can say that $\{\heartsuit\} \Bumpeq \{\clubsuit\}$, for example. As it happens, our new symbol $\Bumpeq$ is an equivalence relation as defined in Definition 15: it's reflexive, because every set has the same cardinality as itself; it's symmetric, because if $|A| = |B|$, then $|B| = |A|$; and it's transitive, because if $|A| = |B|$ and $|B| = |C|$, then $|A| = |C|$.[135]

[133] The fancy term for a set with one element is a *singleton*.

[134] Riddle Y.H.P. this: what is $|\mathcal{P}(\{\heartsuit, \clubsuit, \spadesuit, \diamondsuit\})|$—that is, what's the cardinality of the power set of the set of suits? (Hint: it's not 4.) Now here's a question: can you come up with a rule that tells you the cardinality of the power set of any finite set? (Hint: one common notation for the power set of A is 2^A.)

[135] Notice that $\Bumpeq$ *inherits* these nice properties from = on the natural numbers. A lot of next week will be spent asking: under what conditions do we inherit nice properties from $\geqslant$, along with =?

THAT SAID, THE CARDINALITY OF A SET has its limitations. Notice, for example, that we have

$$|\mathbb{N}| = |\mathbb{R}| = \infty,$$

which just tells us that both $\mathbb{N}$ and $\mathbb{R}$ are infinite sets.[136] But it doesn't tell us anything about how *big* these infinite sets are, and it seems to Y.H.P. that $\mathbb{R}$ is bigger than $\mathbb{N}$. So we need a more refined notion of set size.

[136] For this, you're paying. There are infinite counting numbers! Who knew?!

19 Definition (Cantor's Notion of Size)

Let A and B be sets. We say A and B are the same size *just in case there exists a one-to-one, onto function $f : A \to B$.*

This is a more refined notion of size because it tells us something about how *big* a set is. Notice that, for example, {♥} and {♣} are the same size, because there's a one-to-one, onto function f : {♥} → {♣}; see the table at right. The same goes for a doubleton pair of sets, like {♥, ♣} and {♠, ♦}. Careful inspection of these tables reveals that all we're doing here is changing all the labels on the elements of the domain.[137]

[137] Are you getting a sense of what Y.H.P. means when he says the relationships between things are more important than the things themselves?

$a \in$ {♥}	$f(a) \in$ {♣}
♥	♣

Table 13: A one-to-one, onto function f : {♥} → {♣}.

$a \in$ {♥, ♣}	$f(a) \in$ {♠, ♦}
♥	♠
♣	♦

Table 14: A one-to-one, onto function f : {♥, ♣} → {♠, ♦}.

BUT WHAT ABOUT the cardinality of $\mathbb{N}$ and $\mathbb{R}$? Famously:

20 Theorem (Cantor's Theorem)

There is no one-to-one, onto function $f : \mathbb{N} \to \mathbb{R}$.

This one isn't too hard to prove; you do it with a table. Suppose there existed such a function $f : \mathbb{N} \to \mathbb{R}$. Then we could write out a table like Table 15. Here we've written out the decimal expansion of each real

$n \in \mathbb{N}$	$f(n) \in \mathbb{R}$
1	$0.d_{11}d_{12}d_{13}d_{14}\cdots$
2	$0.d_{21}d_{22}d_{23}d_{24}\cdots$
3	$0.d_{31}d_{32}d_{33}d_{34}\cdots$
4	$0.d_{41}d_{42}d_{43}d_{44}\cdots$
⋮	⋮

Table 15: A one-to-one, onto function $f : \mathbb{N} \to \mathbb{R}$...?

number $f(n)$, where each d_{ij} is a digit. For example, if the number is 0.143, then $d_{11} = 1$, $d_{12} = 4$, and $d_{13} = 3$, and $d_{1j} = 0$ for all $j > 3$. Our goal is to construct a number $x \in \mathbb{R}$ that's not in the table, which would mean that the function f fails to be onto. Well, we can do this by constructing a number x whose ith digit is different from the ith digit of $f(i)$. For each digit d_{ij}, add 1 to it and then take the remainder when dividing by 10—this sends 1 to 2, 2 to 3, and so on, up to 9 to 0. We now have a number that differs from $f(1)$ in the first digit, from $f(2)$ in the second digit, and so on. So it's not in the table, and so f fails to be onto. We conclude that there are more real numbers than natural numbers.[138]

This is called a "diagonal argument." Can you see why? Look which digits we're changing in each $f(n)$. Y.H.P. has even taken the liberty of highlighting the cuprits.

[138] One often refers to the infinity of the natural numbers as *countable infinity*, and the infinity of the real numbers as *uncountable infinity*. There are other infinities, too. In fact, there are infinitely many infinities! But we won't get into that here. Needless to say, Y.H.P. is a big fan of infinity.

Functions are everywhere. Important among these, at least for our purposes, are what are called "utility functions." Utility functions operate on all sorts of domains—that is to say, they can take in different kinds of inputs—but they nearly always return a real number that represents the "utility" or "value" of that input to the agent making the decision. For example, suppose we wanted to model the United States as an independent decision-making agent,[139] and suppose her alternatives were to either attack Canada, threaten to attack Canada, or simply do nothing and maintain the status quo. We might gather these three strategic alternatives into a set,

$$A = \{\text{attack Canada}, \text{threaten to attack Canada}, \text{do nothing}\},$$

and then define a utility function $u : A \to \mathbb{R}$ that assigns a real number to each alternative, representing the United States' preference for that outcome. For instance, we might have

$$\begin{aligned} u(\text{attack Canada}) &= -10, \\ u(\text{threaten to attack Canada}) &= 0, \\ u(\text{do nothing}) &= 5. \end{aligned}$$

Here, higher numbers indicate more preferred outcomes, so that doing nothing is the most preferred outcome, threatening to attack Canada is neutral, and attacking Canada is the least preferred outcome.

What makes a utility function a utility function is the meaning we assign to the numbers it produces: they represent the preferences of the decision-maker over the set of alternatives.[140] We can use utility functions to model states' decisions about wars and threats just as easily as we could your decisions about beverages at the minimart. In either case, we have a function

$$u : A \to \mathbb{R}$$

sending alternatives, however gathered for a given decision-maker, to real numbers representing the decision-maker's preferences over those alternatives.

It should be noted that the story undergirding the utility function model of decision-making is, in a word, dumb. Y.H.P. doesn't know about you, but when he goes into a minimart to get a drink, he doesn't scan the entire freaking cooler assigning a utility value to each possible beverage and then choosing the one with the highest utility.[141] But before we reject it out of hand, we ought to ask ourselves: what would have to be true about a decision-maker's behavior for it to be accurately represented by a utility function? In other words, under what conditions can we say that a decision-maker is acting "as if" they are maximizing a utility function?

[139] Whoa, that's a big assumption! Is the United States the same sort of decision-maker as an individual human being? How would we even begin to assign a utility function to a country? We will discuss the wrinkles associated with such an endeavor in a later section, namely when we turn our attention to the problem of *social choice*.

[140] Again, it is incumbent upon us to imbue these things with meaning; the numbers themselves are just numbers until we interpret them as representing preferences. And just as importantly: the machine that produces the numbers (the utility function) is itself a mathematical object that we define; it does not have intrinsic meaning outside of the interpretation we give it. To imbue a function with meaning is to assert some sort of position about the nature of a *relationship*.
Just fair warning: once we toggle over into rational choice theory in a second, we'll often use X to denote alternatives, not A. The symbols, as ever, remain arbitrary.

[141] He just grabs that kombucha, baby, and he never lets it go.

And this is the beginning of the story of *rational choice theory*, which seeks to formalize the conditions under which we can model decision-makers as utility maximizers. And that is the subject of the next set of notes.

References