

Part I
Methods

This part covers general and methodological material, especially explanations of various proof methods a non-mathematics student may be unfamiliar with. It currently contains a chapter on how to write proofs, and a chapter on induction, but additional sections for those, exercises, and a chapter on mathematical terminology is also planned.

Chapter 1

Proofs

1.1 Introduction

math:prf:int:
sec

Based on your experiences in introductory logic, you might be comfortable with a **derivation** system—probably a natural deduction or Fitch style **derivation** system, or perhaps a proof-tree system. You probably remember doing proofs in these systems, either proving a **formula** or show that a given argument is valid. In order to do this, you applied the rules of the system until you got the desired end result. In reasoning *about* logic, we also prove things, but in most cases we are not using a **derivation** system. In fact, most of the proofs we consider are done in English (perhaps, with some symbolic language thrown in) rather than entirely in the language of first-order logic. When constructing such proofs, you might at first be at a loss—how do I prove something without a **derivation** system? How do I start? How do I know if my proof is correct?

Before attempting a proof, it's important to know what a proof is and how to construct one. As implied by the name, a *proof* is meant to show that something is true. You might think of this in terms of a dialogue—someone asks you if something is true, say, if every prime other than two is an odd number. To answer “yes” is not enough; they might want to know *why*. In this case, you'd give them a proof.

In everyday discourse, it might be enough to gesture at an answer, or give an incomplete answer. In logic and mathematics, however, we want rigorous proof—we want to show that something is true beyond *any* doubt. This means that every step in our proof must be justified, and the justification must be cogent (i.e., the assumption you're using is actually assumed in the statement of the theorem you're proving, the definitions you apply must be correctly applied, the justifications appealed to must be correct inferences, etc.).

Usually, we're proving some statement. We call the statements we're proving by various names: propositions, theorems, lemmas, or corollaries. A proposition is a basic proof-worthy statement: important enough to record, but perhaps not particularly deep nor applied often. A theorem is a significant, important proposition. Its proof often is broken into several steps, and some-

times it is named after the person who first proved it (e.g., Cantor’s Theorem, the Löwenheim–Skolem theorem) or after the fact it concerns (e.g., the completeness theorem). A lemma is a proposition or theorem that is used in the proof of a more important result. Confusingly, sometimes lemmas are important results in themselves, and also named after the person who introduced them (e.g., Zorn’s Lemma). A corollary is a result that easily follows from another one.

A statement to be proved often contains assumptions that clarify which kinds of things we’re proving something about. It might begin with “Let φ be a formula of the form $\psi \rightarrow \chi$ ” or “Suppose $\Gamma \vdash \varphi$ ” or something of the sort. These are *hypotheses* of the proposition, theorem, or lemma, and you may assume these to be true in your proof. They restrict what we’re proving, and also introduce some names for the objects we’re talking about. For instance, if your proposition begins with “Let φ be a formula of the form $\psi \rightarrow \chi$,” you’re proving something about all formulas of a certain sort only (namely, conditionals), and it’s understood that $\psi \rightarrow \chi$ is an arbitrary conditional that your proof will talk about.

1.2 Starting a Proof

But where do you even start?

You’ve been given something to prove, so this should be the last thing that is mentioned in the proof (you can, obviously, *announce* that you’re going to prove it at the beginning, but you don’t want to use it as an assumption). Write what you are trying to prove at the bottom of a fresh sheet of paper—this way you don’t lose sight of your goal.

Next, you may have some assumptions that you are able to use (this will be made clearer when we talk about the *type* of proof you are doing in the next section). Write these at the top of the page and make sure to flag that they are assumptions (i.e., if you are assuming p , write “assume that p ,” or “suppose that p ”). Finally, there might be some definitions in the question that you need to know. You might be told to use a specific definition, or there might be various definitions in the assumptions or conclusion that you are working towards. *Write these down and ensure that you understand what they mean.*

How you set up your proof will also be dependent upon the form of the question. The next section provides details on how to set up your proof based on the type of sentence.

1.3 Using Definitions

We mentioned that you must be familiar with all definitions that may be used in the proof, and that you can properly apply them. This is a really important point, and it is worth looking at in a bit more detail. Definitions are used to abbreviate properties and relations so we can talk about them more succinctly. The introduced abbreviation is called the *definiendum*, and what it abbreviates

is the *definiens*. In proofs, we often have to go back to how the definiendum was introduced, because we have to exploit the logical structure of the definiens (the long version of which the defined term is the abbreviation) to get through our proof. By unpacking definitions, you’re ensuring that you’re getting to the heart of where the logical action is.

We’ll start with an example. Suppose you want to prove the following:

Proposition 1.1. *For any sets A and B , $A \cup B = B \cup A$.*

In order to even start the proof, we need to know what it means for two sets to be identical; i.e., we need to know what the “=” in that equation means for sets. Sets are defined to be identical whenever they have the same **elements**. So the definition we have to unpack is:

Definition 1.2. Sets A and B are *identical*, $A = B$, iff every **element** of A is **an element** of B , and vice versa.

This definition uses A and B as placeholders for arbitrary sets. What it defines—the *definiendum*—is the expression “ $A = B$ ” by giving the condition under which $A = B$ is true. This condition—“every **element** of A is **an element** of B , and vice versa”—is the *definiens*.¹ The definition specifies that $A = B$ is true if, and only if (we abbreviate this to “iff”) the condition holds.

When you apply the definition, you have to match the A and B in the definition to the case you’re dealing with. In our case, it means that in order for $A \cup B = B \cup A$ to be true, each $z \in A \cup B$ must also be in $B \cup A$, and vice versa. The expression $A \cup B$ in the proposition plays the role of A in the definition, and $B \cup A$ that of B . Since A and B are used both in the definition and in the statement of the proposition we’re proving, but in different uses, you have to be careful to make sure you don’t mix up the two. For instance, it would be a mistake to think that you could prove the proposition by showing that every **element** of A is **an element** of B , and vice versa—that would show that $A = B$, not that $A \cup B = B \cup A$. (Also, since A and B may be any two sets, you won’t get very far, because if nothing is assumed about A and B they may well be different sets.)

Within the proof we are dealing with set-theoretic notions such as union, and so we must also know the meanings of the symbol \cup in order to understand how the proof should proceed. And sometimes, unpacking the definition gives rise to further definitions to unpack. For instance, $A \cup B$ is defined as $\{z : z \in A \text{ or } z \in B\}$. So if you want to prove that $x \in A \cup B$, unpacking the definition of \cup tells you that you have to prove $x \in \{z : z \in A \text{ or } z \in B\}$. Now you also have to remember that $x \in \{z : \dots z \dots\}$ iff $\dots x \dots$. So, further unpacking the definition of the $\{z : \dots z \dots\}$ notation, what you have to show is: $x \in A$ or $x \in B$. So, “every **element** of $A \cup B$ is also **an element** of $B \cup A$ ” really means:

¹In this particular case—and very confusingly!—when $A = B$, the sets A and B are just one and the same set, even though we use different letters for it on the left and the right side. But the ways in which that set is picked out may be different, and that makes the definition non-trivial.

“for every x , if $x \in A$ or $x \in B$, then $x \in B$ or $x \in A$.” If we fully unpack the definitions in the proposition, we see that what we have to show is this:

Proposition 1.3. *For any sets A and B : (a) for every x , if $x \in A$ or $x \in B$, then $x \in B$ or $x \in A$, and (b) for every x , if $x \in B$ or $x \in A$, then $x \in A$ or $x \in B$.*

What’s important is that unpacking definitions is a necessary part of constructing a proof. Properly doing it is sometimes difficult: you must be careful to distinguish and match the variables in the definition and the terms in the claim you’re proving. In order to be successful, you must know what the question is asking and what all the terms used in the question mean—you will often need to unpack more than one definition. In simple proofs such as the ones below, the solution follows almost immediately from the definitions themselves. Of course, it won’t always be this simple.

Problem 1.1. Suppose you are asked to prove that $A \cap B \neq \emptyset$. Unpack all the definitions occurring here, i.e., restate this in a way that does not mention “ \cap ”, “ $=$ ”, or “ \emptyset ”.

1.4 Inference Patterns

Proofs are composed of individual inferences. When we make an inference, we typically indicate that by using a word like “so,” “thus,” or “therefore.” The inference often relies on one or two facts we already have available in our proof—it may be something we have assumed, or something that we’ve concluded by an inference already. To be clear, we may label these things, and in the inference we indicate what other statements we’re using in the inference. An inference will often also contain an explanation of *why* our new conclusion follows from the things that come before it. There are some common patterns of inference that are used very often in proofs; we’ll go through some below. Some patterns of inference, like proofs by induction, are more involved (and will be discussed later).

[math:prf:pat:
sec](#)

We’ve already discussed one pattern of inference: unpacking, or applying, a definition. When we unpack a definition, we just restate something that involves the definiendum by using the definiens. For instance, suppose that we have already established in the course of a proof that $D = E$ (a). Then we may apply the definition of $=$ for sets and infer: “Thus, by definition from (a), every **element** of D is an **element** of E and vice versa.”

Somewhat confusingly, we often do not write the justification of an inference when we actually make it, but before. Suppose we haven’t already proved that $D = E$, but we want to. If $D = E$ is the conclusion we aim for, then we can restate this aim also by applying the definition: to prove $D = E$ we have to prove that every **element** of D is an **element** of E and vice versa. So our proof will have the form: (a) prove that every **element** of D is an **element** of E ; (b) every **element** of E is an **element** of D ; (c) therefore, from (a) and (b) by

definition of $=$, $D = E$. But we would usually not write it this way. Instead we might write something like,

We want to show $D = E$. By definition of $=$, this amounts to showing that every element of D is an element of E and vice versa.

(a) ... (a proof that every element of D is an element of E) ...

(b) ... (a proof that every element of E is an element of D) ...

Using a Conjunction

Perhaps the simplest inference pattern is that of drawing as conclusion one of the conjuncts of a conjunction. In other words: if we have assumed or already proved that p and q , then we're entitled to infer that p (and also that q). This is such a basic inference that it is often not mentioned. For instance, once we've unpacked the definition of $D = E$ we've established that every element of D is an element of E and vice versa. From this we can conclude that every element of E is an element of D (that's the "vice versa" part).

Proving a Conjunction

Sometimes what you'll be asked to prove will have the form of a conjunction; you will be asked to "prove p and q ." In this case, you simply have to do two things: prove p , and then prove q . You could divide your proof into two sections, and for clarity, label them. When you're making your first notes, you might write "(1) Prove p " at the top of the page, and "(2) Prove q " in the middle of the page. (Of course, you might not be explicitly asked to prove a conjunction but find that your proof requires that you prove a conjunction. For instance, if you're asked to prove that $D = E$ you will find that, after unpacking the definition of $=$, you have to prove: every element of D is an element of E and every element of E is an element of D).

Proving a Disjunction

When what you are proving takes the form of a disjunction (i.e., it is a statement of the form " p or q "), it is enough to show that one of the disjuncts is true. However, it basically never happens that either disjunct just follows from the assumptions of your theorem. More often, the assumptions of your theorem are themselves disjunctive, or you're showing that all things of a certain kind have one of two properties, but some of the things have the one and others have the other property. This is where proof by cases is useful (see below).

Conditional Proof

Many theorems you will encounter are in conditional form (i.e., show that if p holds, then q is also true). These cases are nice and easy to set up—simply assume the antecedent of the conditional (in this case, p) and prove the

conclusion q from it. So if your theorem reads, “If p then q ,” you start your proof with “assume p ” and at the end you should have proved q .

Conditionals may be stated in different ways. So instead of “If p then q ,” a theorem may state that “ p only if q ,” “ q if p ,” or “ q , provided p .” These all mean the same and require assuming p and proving q from that assumption. Recall that a biconditional (“ p if and only if (iff) q ”) is really two conditionals put together: if p then q , and if q then p . All you have to do, then, is two instances of conditional proof: one for the first conditional and another one for the second. Sometimes, however, it is possible to prove an “iff” statement by chaining together a bunch of other “iff” statements so that you start with “ p ” an end with “ q ”—but in that case you have to make sure that each step really is an “iff.”

Universal Claims

Using a universal claim is simple: if something is true for anything, it’s true for each particular thing. So if, say, the hypothesis of your proof is $A \subseteq B$, that means (unpacking the definition of \subseteq), that, for every $x \in A$, $x \in B$. Thus, if you already know that $z \in A$, you can conclude $z \in B$.

Proving a universal claim may seem a little bit tricky. Usually these statements take the following form: “If x has P , then it has Q ” or “All P s are Q s.” Of course, it might not fit this form perfectly, and it takes a bit of practice to figure out what you’re asked to prove exactly. But: we often have to prove that all objects with some property have a certain other property.

The way to prove a universal claim is to introduce names or variables, for the things that have the one property and then show that they also have the other property. We might put this by saying that to prove something for *all* P s you have to prove it for an *arbitrary* P . And the name introduced is a name for an arbitrary P . We typically use single letters as these names for arbitrary things, and the letters usually follow conventions: e.g., we use n for natural numbers, φ for **formulas**, A for sets, f for functions, etc.

The trick is to maintain generality throughout the proof. You start by assuming that an arbitrary object (“ x ”) has the property P , and show (based only on definitions or what you are allowed to assume) that x has the property Q . Because you have not stipulated what x is specifically, other than that it has the property P , then you can assert that all every P has the property Q . In short, x is a stand-in for *all* things with property P .

Proposition 1.4. *For all sets A and B , $A \subseteq A \cup B$.*

Proof. Let A and B be arbitrary sets. We want to show that $A \subseteq A \cup B$. By definition of \subseteq , this amounts to: for every x , if $x \in A$ then $x \in A \cup B$. So let $x \in A$ be an arbitrary **element** of A . We have to show that $x \in A \cup B$. Since $x \in A$, $x \in A$ or $x \in B$. Thus, $x \in \{x : x \in A \vee x \in B\}$. But that, by definition of \cup , means $x \in A \cup B$. \square

Proof by Cases

Suppose you have a disjunction as an assumption or as an already established conclusion—you have assumed or proved that p or q is true. You want to prove r . You do this in two steps: first you assume that p is true, and prove r , then you assume that q is true and prove r again. This works because we assume or know that one of the two alternatives holds. The two steps establish that either one is sufficient for the truth of r . (If both are true, we have not one but two reasons for why r is true. It is not necessary to separately prove that r is true assuming both p and q .) To indicate what we’re doing, we announce that we “distinguish cases.” For instance, suppose we know that $x \in B \cup C$. $B \cup C$ is defined as $\{x : x \in B \text{ or } x \in C\}$. In other words, by definition, $x \in B$ or $x \in C$. We would prove that $x \in A$ from this by first assuming that $x \in B$, and proving $x \in A$ from this assumption, and then assume $x \in C$, and again prove $x \in A$ from this. You would write “We distinguish cases” under the assumption, then “Case (1): $x \in B$ ” underneath, and “Case (2): $x \in C$ ” halfway down the page. Then you’d proceed to fill in the top half and the bottom half of the page.

Proof by cases is especially useful if what you’re proving is itself disjunctive. Here’s a simple example:

Proposition 1.5. *Suppose $B \subseteq D$ and $C \subseteq E$. Then $B \cup C \subseteq D \cup E$.*

Proof. Assume (a) that $B \subseteq D$ and (b) $C \subseteq E$. By definition, any $x \in B$ is also $\in D$ (c) and any $x \in C$ is also $\in E$ (d). To show that $B \cup C \subseteq D \cup E$, we have to show that if $x \in B \cup C$ then $x \in D \cup E$ (by definition of \subseteq). $x \in B \cup C$ iff $x \in B$ or $x \in C$ (by definition of \cup). Similarly, $x \in D \cup E$ iff $x \in D$ or $x \in E$. So, we have to show: for any x , if $x \in B$ or $x \in C$, then $x \in D$ or $x \in E$.

So far we’ve only unpacked definitions! We’ve reformulated our proposition without \subseteq and \cup and are left with trying to prove a universal conditional claim. By what we’ve discussed above, this is done by assuming that x is something about which we assume the “if” part is true, and we’ll go on to show that the “then” part is true as well. In other words, we’ll assume that $x \in B$ or $x \in C$ and show that $x \in D$ or $x \in E$.²

Suppose that $x \in B$ or $x \in C$. We have to show that $x \in D$ or $x \in E$. We distinguish cases.

Case 1: $x \in B$. By (c), $x \in D$. Thus, $x \in D$ or $x \in E$. (Here we’ve made the inference discussed in the preceding subsection!)

Case 2: $x \in C$. By (d), $x \in E$. Thus, $x \in D$ or $x \in E$. □

²This paragraph just explains what we’re doing—it’s not part of the proof, and you don’t have to go into all this detail when you write down your own proofs.

Proving an Existence Claim

When asked to prove an existence claim, the question will usually be of the form “prove that there is an x such that $\dots x \dots$ ”, i.e., that some object that has the property described by “ $\dots x \dots$ ”. In this case you’ll have to identify a suitable object show that it has the required property. This sounds straightforward, but a proof of this kind can be tricky. Typically it involves *constructing* or *defining* an object and proving that the object so defined has the required property. Finding the right object may be hard, proving that it has the required property may be hard, and sometimes it’s even tricky to show that you’ve succeeded in defining an object at all!

Generally, you’d write this out by specifying the object, e.g., “let x be \dots ” (where \dots specifies which object you have in mind), possibly proving that \dots in fact describes an object that exists, and then go on to show that x has the property Q . Here’s a simple example.

Proposition 1.6. *Suppose that $x \in B$. Then there is an A such that $A \subseteq B$ and $A \neq \emptyset$.*

Proof. Assume $x \in B$. Let $A = \{x\}$.

Here we’ve defined the set A by enumerating its **elements**. Since we assume that x is an object, and we can always form a set by enumerating its **elements**, we don’t have to show that we’ve succeeded in defining a set A here. However, we still have to show that A has the properties required by the proposition. The proof isn’t complete without that!

Since $x \in A$, $A \neq \emptyset$.

This relies on the definition of A as $\{x\}$ and the obvious facts that $x \in \{x\}$ and $x \notin \emptyset$.

Since x is the only **element** of $\{x\}$, and $x \in B$, every **element** of A is also an **element** of B . By definition of \subseteq , $A \subseteq B$. \square

Using Existence Claims

Suppose you know that some existence claim is true (you’ve proved it, or it’s a hypothesis you can use), say, “for some x , $x \in A$ ” or “there is an $x \in A$.” If you want to use it in your proof, you can just pretend that you have a name for one of the things which your hypothesis says exist. Since A contains at least one thing, there are things to which that name might refer. You might of course not be able to pick one out or describe it further (other than that it is $\in A$). But for the purpose of the proof, you can pretend that you have picked it out and give a name to it. It’s important to pick a name that you haven’t already used (or that appears in your hypotheses), otherwise things can go wrong. In your proof, you indicate this by going from “for some x , $x \in A$ ” to

“Let $a \in A$.” Now you can reason about a , use some other hypotheses, etc., until you come to a conclusion, p . If p no longer mentions a , p is independent of the assumption that $a \in A$, and you’ve shown that it follows just from the assumption “for some x , $x \in A$.”

Proposition 1.7. *If $A \neq \emptyset$, then $A \cup B \neq \emptyset$.*

Proof. Suppose $A \neq \emptyset$. So for some x , $x \in A$.

Here we first just restated the hypothesis of the proposition. This hypothesis, i.e., $A \neq \emptyset$, hides an existential claim, which you get to only by unpacking a few definitions. The definition of $=$ tells us that $A = \emptyset$ iff every $x \in A$ is also $\in \emptyset$ and every $x \in \emptyset$ is also $\in A$. Negating both sides, we get: $A \neq \emptyset$ iff either some $x \in A$ is $\notin \emptyset$ or some $x \in \emptyset$ is $\notin A$. Since nothing is $\in \emptyset$, the second disjunct can never be true, and “ $x \in A$ and $x \notin \emptyset$ ” reduces to just $x \in A$. So $A \neq \emptyset$ iff for some x , $x \in A$. That’s an existence claim. Now we use that existence claim by introducing a name for one of the **elements** of A :

Let $a \in A$.

Now we’ve introduced a name for one of the things $\in A$. We’ll continue to argue about a , but we’ll be careful to only assume that $a \in A$ and nothing else:

Since $a \in A$, $a \in A \cup B$, by definition of \cup . So for some x , $x \in A \cup B$, i.e., $A \cup B \neq \emptyset$.

In that last step, we went from “ $a \in A \cup B$ ” to “for some x , $x \in A \cup B$.” That doesn’t mention a anymore, so we know that “for some x , $x \in A \cup B$ ” follows from “for some x , $x \in A$ alone.” But that means that $A \cup B \neq \emptyset$. □

It’s maybe good practice to keep bound variables like “ x ” separate from hypothetical names like a , like we did. In practice, however, we often don’t and just use x , like so:

Suppose $A \neq \emptyset$, i.e., there is an $x \in A$. By definition of \cup , $x \in A \cup B$.
So $A \cup B \neq \emptyset$.

However, when you do this, you have to be extra careful that you use different x ’s and y ’s for different existential claims. For instance, the following is *not* a correct proof of “If $A \neq \emptyset$ and $B \neq \emptyset$ then $A \cap B \neq \emptyset$ ” (which is not true).

Suppose $A \neq \emptyset$ and $B \neq \emptyset$. So for some x , $x \in A$ and also for some x , $x \in B$. Since $x \in A$ and $x \in B$, $x \in A \cap B$, by definition of \cap .
So $A \cap B \neq \emptyset$.

Can you spot where the incorrect step occurs and explain why the result does not hold?

1.5 An Example

Our first example is the following simple fact about unions and intersections of sets. It will illustrate unpacking definitions, proofs of conjunctions, of universal claims, and proof by cases. mth:prf:ex1:
sec

Proposition 1.8. *For any sets A , B , and C , $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$*

Let's prove it!

Proof. We want to show that for any sets A , B , and C , $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

First we unpack the definition of “=” in the statement of the proposition. Recall that proving sets identical means showing that the sets have the same elements. That is, all elements of $A \cup (B \cap C)$ are also elements of $(A \cup B) \cap (A \cup C)$, and vice versa. The “vice versa” means that also every element of $(A \cup B) \cap (A \cup C)$ must be an element of $A \cup (B \cap C)$. So in unpacking the definition, we see that we have to prove a conjunction. Let's record this:

By definition, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ iff every element of $A \cup (B \cap C)$ is also an element of $(A \cup B) \cap (A \cup C)$, and every element of $(A \cup B) \cap (A \cup C)$ is an element of $A \cup (B \cap C)$.

Since this is a conjunction, we must prove each conjunct separately. Lets start with the first: let's prove that every element of $A \cup (B \cap C)$ is also an element of $(A \cup B) \cap (A \cup C)$.

This is a universal claim, and so we consider an arbitrary element of $A \cup (B \cap C)$ and show that it must also be an element of $(A \cup B) \cap (A \cup C)$. We'll pick a variable to call this arbitrary element by, say, z . Our proof continues:

First, we prove that every element of $A \cup (B \cap C)$ is also an element of $(A \cup B) \cap (A \cup C)$. Let $z \in A \cup (B \cap C)$. We have to show that $z \in (A \cup B) \cap (A \cup C)$.

Now it is time to unpack the definition of \cup and \cap . For instance, the definition of \cup is: $A \cup B = \{z : z \in A \text{ or } z \in B\}$. When we apply the definition to “ $A \cup (B \cap C)$,” the role of the “ B ” in the definition is now played by “ $B \cap C$,” so $A \cup (B \cap C) = \{z : z \in A \text{ or } z \in B \cap C\}$. So our assumption that $z \in A \cup (B \cap C)$ amounts to: $z \in \{z : z \in A \text{ or } z \in B \cap C\}$. And $z \in \{z : \dots z \dots\}$ iff $\dots z \dots$, i.e., in this case, $z \in A$ or $z \in B \cap C$.

By the definition of \cup , either $z \in A$ or $z \in B \cap C$.

Since this is a disjunction, it will be useful to apply proof by cases. We take the two cases, and show that in each one, the conclusion we're aiming for (namely, “ $z \in (A \cup B) \cap (A \cup C)$ ”) obtains.

Case 1: Suppose that $z \in A$.

There's not much more to work from based on our assumptions. So let's look at what we have to work with in the conclusion. We want to show that $z \in (A \cup B) \cap (A \cup C)$. Based on the definition of \cap , if we want to show that $z \in (A \cup B) \cap (A \cup C)$, we have to show that it's in both $(A \cup B)$ and $(A \cup C)$. But $z \in A \cup B$ iff $z \in A$ or $z \in B$, and we already have (as the assumption of case 1) that $z \in A$. By the same reasoning—switching C for B — $z \in A \cup C$. This argument went in the reverse direction, so let's record our reasoning in the direction needed in our proof.

Since $z \in A$, $z \in A$ or $z \in B$, and hence, by definition of \cup , $z \in A \cup B$. Similarly, $z \in A \cup C$. But this means that $z \in (A \cup B) \cap (A \cup C)$, by definition of \cap .

This completes the first case of the proof by cases. Now we want to derive the conclusion in the second case, where $z \in B \cap C$.

Case 2: Suppose that $z \in B \cap C$.

Again, we are working with the intersection of two sets. Let's apply the definition of \cap :

Since $z \in B \cap C$, z must be **an element** of both B and C , by definition of \cap .

It's time to look at our conclusion again. We have to show that z is in both $(A \cup B)$ and $(A \cup C)$. And again, the solution is immediate.

Since $z \in B$, $z \in (A \cup B)$. Since $z \in C$, also $z \in (A \cup C)$. So, $z \in (A \cup B) \cap (A \cup C)$.

Here we applied the definitions of \cup and \cap again, but since we've already recalled those definitions, and already showed that if z is in one of two sets it is in their union, we don't have to be as explicit in what we've done.

We've completed the second case of the proof by cases, so now we can assert our first conclusion.

So, if $z \in A \cup (B \cap C)$ then $z \in (A \cup B) \cap (A \cup C)$.

Now we just want to show the other direction, that every **element** of $(A \cup B) \cap (A \cup C)$ is **an element** of $A \cup (B \cap C)$. As before, we prove this universal claim by assuming we have an arbitrary **element** of the first set and show it must be in the second set. Let's state what we're about to do.

Now, assume that $z \in (A \cup B) \cap (A \cup C)$. We want to show that $z \in A \cup (B \cap C)$.

We are now working from the hypothesis that $z \in (A \cup B) \cap (A \cup C)$. It hopefully isn't too confusing that we're using the same z here as in the first part of the proof. When we finished that part, all the assumptions we've made there are no longer in effect, so now we can make new assumptions about what z is. If that is confusing to you, just replace z with a different variable in what follows.

We know that z is in both $A \cup B$ and $A \cup C$, by definition of \cap . And by the definition of \cup , we can further unpack this to: either $z \in A$ or $z \in B$, and also either $z \in A$ or $z \in C$. This looks like a proof by cases again—except the “and” makes it confusing. You might think that this amounts to there being three possibilities: z is either in A , B or C . But that would be a mistake. We have to be careful, so let's consider each disjunction in turn.

By definition of \cap , $z \in A \cup B$ and $z \in A \cup C$. By definition of \cup , $z \in A$ or $z \in B$. We distinguish cases.

Since we're focusing on the first disjunction, we haven't gotten our second disjunction (from unpacking $A \cup C$) yet. In fact, we don't need it yet. The first case is $z \in A$, and **an element** of a set is also **an element** of the union of that set with any other. So case 1 is easy:

Case 1: Suppose that $z \in A$. It follows that $z \in A \cup (B \cap C)$.

Now for the second case, $z \in B$. Here we'll unpack the second \cup and do another proof-by-cases:

Case 2: Suppose that $z \in B$. Since $z \in A \cup C$, either $z \in A$ or $z \in C$. We distinguish cases further:

Case 2a: $z \in A$. Then, again, $z \in A \cup (B \cap C)$.

Ok, this was a bit weird. We didn't actually need the assumption that $z \in B$ for this case, but that's ok.

Case 2b: $z \in C$. Then $z \in B$ and $z \in C$, so $z \in B \cap C$, and consequently, $z \in A \cup (B \cap C)$.

This concludes both proofs-by-cases and so we're done with the second half.

So, if $z \in (A \cup B) \cap (A \cup C)$ then $z \in A \cup (B \cap C)$. □

1.6 Another Example

Proposition 1.9. *If $A \subseteq C$, then $A \cup (C \setminus A) = C$.*

meth:prfex2:
sec

Proof. Suppose that $A \subseteq C$. We want to show that $A \cup (C \setminus A) = C$.

We begin by observing that this is a conditional statement. It is tacitly universally quantified: the proposition holds for all sets A and C . So A and C are variables for arbitrary sets. To prove such a statement, we assume the antecedent and prove the consequent.

We continue by using the assumption that $A \subseteq C$. Let's unpack the definition of \subseteq : the assumption means that all **elements** of A are also **elements** of C . Let's write this down—it's an important fact that we'll use throughout the proof.

By the definition of \subseteq , since $A \subseteq C$, for all z , if $z \in A$, then $z \in C$.

We've unpacked all the definitions that are given to us in the assumption. Now we can move onto the conclusion. We want to show that $A \cup (C \setminus A) = C$, and so we set up a proof similarly to the last example: we show that every **element** of $A \cup (C \setminus A)$ is also **an element** of C and, conversely, every **element** of C is **an element** of $A \cup (C \setminus A)$. We can shorten this to: $A \cup (C \setminus A) \subseteq C$ and $C \subseteq A \cup (C \setminus A)$. (Here we're doing the opposite of unpacking a definition, but it makes the proof a bit easier to read.) Since this is a conjunction, we have to prove both parts. To show the first part, i.e., that every **element** of $A \cup (C \setminus A)$ is also **an element** of C , we assume that $z \in A \cup (C \setminus A)$ for an arbitrary z and show that $z \in C$. By the definition of \cup , we can conclude that $z \in A$ or $z \in C \setminus A$ from $z \in A \cup (C \setminus A)$. You should now be getting the hang of this.

$A \cup (C \setminus A) = C$ iff $A \cup (C \setminus A) \subseteq C$ and $C \subseteq (A \cup (C \setminus A))$. First we prove that $A \cup (C \setminus A) \subseteq C$. Let $z \in A \cup (C \setminus A)$. So, either $z \in A$ or $z \in (C \setminus A)$.

We've arrived at a disjunction, and from it we want to prove that $z \in C$. We do this using proof by cases.

Case 1: $z \in A$. Since for all z , if $z \in A$, $z \in C$, we have that $z \in C$.

Here we've used the fact recorded earlier which followed from the hypothesis of the proposition that $A \subseteq C$. The first case is complete, and we turn to the second case, $z \in (C \setminus A)$. Recall that $C \setminus A$ denotes the *difference* of the two sets, i.e., the set of all **elements** of C which are not **elements** of A . But any **element** of C not in A is in particular **an element** of C .

Case 2: $z \in (C \setminus A)$. This means that $z \in C$ and $z \notin A$. So, in particular, $z \in C$.

Great, we've proved the first direction. Now for the second direction. Here we prove that $C \subseteq A \cup (C \setminus A)$. So we assume that $z \in C$ and prove that $z \in A \cup (C \setminus A)$.

Now let $z \in C$. We want to show that $z \in A$ or $z \in C \setminus A$.

Since all **elements** of A are also **elements** of C , and $C \setminus A$ is the set of all things that are **elements** of C but not A , it follows that z is either in A or in $C \setminus A$. This may be a bit unclear if you don't already know why the result is true. It would be better to prove it step-by-step. It will help to use a simple fact which we can state without proof: $z \in A$ or $z \notin A$. This is called the “principle of excluded middle:” for any statement p , either p is true or its negation is true. (Here, p is the statement that $z \in A$.) Since this is a disjunction, we can again use proof-by-cases.

Either $z \in A$ or $z \notin A$. In the former case, $z \in A \cup (C \setminus A)$. In the latter case, $z \in C$ and $z \notin A$, so $z \in C \setminus A$. But then $z \in A \cup (C \setminus A)$.

Our proof is complete: we have shown that $A \cup (C \setminus A) = C$. □

1.7 Proof by Contradiction

In the first instance, proof by contradiction is an inference pattern that is used to prove negative claims. Suppose you want to show that some claim p is *false*, i.e., you want to show $\neg p$. The most promising strategy is to (a) suppose that p is true, and (b) show that this assumption leads to something you know to be false. “Something known to be false” may be a result that conflicts with—contradicts— p itself, or some other hypothesis of the overall claim you are considering. For instance, a proof of “if q then $\neg p$ ” involves assuming that q is true and proving $\neg p$ from it. If you prove $\neg p$ by contradiction, that means assuming p in addition to q . If you can prove $\neg q$ from p , you have shown that the assumption p leads to something that contradicts your other assumption q , since q and $\neg q$ cannot both be true. Of course, you have to use other inference patterns in your proof of the contradiction, as well as unpacking definitions. Let's consider an example.

[mth:prf:con:
sec](#)

Proposition 1.10. *If $A \subseteq B$ and $B = \emptyset$, then A has no **elements**.*

Proof. Suppose $A \subseteq B$ and $B = \emptyset$. We want to show that A has no **elements**.

Since this is a conditional claim, we assume the antecedent and want to prove the consequent. The consequent is: A has no **elements**. We can make that a bit more explicit: it's not the case that there is an $x \in A$.

A has no **elements** iff it's not the case that there is an x such that $x \in A$.

So we've determined that what we want to prove is really a negative claim $\neg p$, namely: it's not the case that there is an $x \in A$. To use proof by contradiction, we have to assume the corresponding positive claim p , i.e., there is an $x \in A$, and prove a contradiction from it. We indicate that we're doing a proof by contradiction by

writing “by way of contradiction, assume” or even just “suppose not,” and then state the assumption p .

Suppose not: there is an $x \in A$.

This is now the new assumption we’ll use to obtain a contradiction. We have two more assumptions: that $A \subseteq B$ and that $B = \emptyset$. The first gives us that $x \in B$:

Since $A \subseteq B$, $x \in B$.

But since $B = \emptyset$, every **element** of B (e.g., x) must also be an **element** of \emptyset .

Since $B = \emptyset$, $x \in \emptyset$. This is a contradiction, since by definition \emptyset has no **elements**.

This already completes the proof: we’ve arrived at what we need (a contradiction) from the assumptions we’ve set up, and this means that the assumptions can’t all be true. Since the first two assumptions ($A \subseteq B$ and $B = \emptyset$) are not contested, it must be the last assumption introduced (there is an $x \in A$) that must be false. But if we want to be thorough, we can spell this out.

Thus, our assumption that there is an $x \in A$ must be false, hence, A has no **elements** by proof by contradiction. \square

Every positive claim is trivially equivalent to a negative claim: p iff $\neg\neg p$. So proofs by contradiction can also be used to establish positive claims “indirectly,” as follows: To prove p , read it as the negative claim $\neg\neg p$. If we can prove a contradiction from $\neg p$, we’ve established $\neg\neg p$ by proof by contradiction, and hence p .

In the last example, we aimed to prove a negative claim, namely that A has no **elements**, and so the assumption we made for the purpose of proof by contradiction (i.e., that there is an $x \in A$) was a positive claim. It gave us something to work with, namely the hypothetical $x \in A$ about which we continued to reason until we got to $x \in \emptyset$.

When proving a positive claim indirectly, the assumption you’d make for the purpose of proof by contradiction would be negative. But very often you can easily reformulate a positive claim as a negative claim, and a negative claim as a positive claim. Our previous proof would have been essentially the same had we proved “ $A = \emptyset$ ” instead of the negative consequent “ A has no **elements**.” (By definition of $=$, “ $A = \emptyset$ ” is a general claim, since it unpacks to “every **element** of A is an **element** of \emptyset and vice versa”.) But it is easily seen to be equivalent to the negative claim “not: there is an $x \in A$.”

So it is sometimes easier to work with $\neg p$ as an assumption than it is to prove p directly. Even when a direct proof is just as simple or even simpler (as in the next examples), some people prefer to proceed indirectly. If the double

negation confuses you, think of a proof by contradiction of some claim as a proof of a contradiction from the *opposite* claim. So, a proof by contradiction of $\neg p$ is a proof of a contradiction from the assumption p ; and proof by contradiction of p is a proof of a contradiction from $\neg p$.

Proposition 1.11. $A \subseteq A \cup B$.

Proof. We want to show that $A \subseteq A \cup B$.

On the face of it, this is a positive claim: every $x \in A$ is also in $A \cup B$. The negation of that is: some $x \in A$ is $\notin A \cup B$. So we can prove the claim indirectly by assuming this negated claim, and showing that it leads to a contradiction.

Suppose not, i.e., $A \not\subseteq A \cup B$.

We have a definition of $A \subseteq A \cup B$: every $x \in A$ is also $\in A \cup B$. To understand what $A \not\subseteq A \cup B$ means, we have to use some elementary logical manipulation on the unpacked definition: it's false that every $x \in A$ is also $\in A \cup B$ iff there is *some* $x \in A$ that is $\notin C$. (This is a place where you want to be very careful: many students' attempted proofs by contradiction fail because they analyze the negation of a claim like "all As are Bs" incorrectly.) In other words, $A \not\subseteq A \cup B$ iff there is an x such that $x \in A$ and $x \notin A \cup B$. From then on, it's easy.

So, there is an $x \in A$ such that $x \notin A \cup B$. By definition of \cup , $x \in A \cup B$ iff $x \in A$ or $x \in B$. Since $x \in A$, we have $x \in A \cup B$. This contradicts the assumption that $x \notin A \cup B$. \square

Problem 1.2. Prove *indirectly* that $A \cap B \subseteq A$.

Proposition 1.12. If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Proof. Suppose $A \subseteq B$ and $B \subseteq C$. We want to show $A \subseteq C$.

Let's proceed indirectly: we assume the negation of what we want to establish.

Suppose not, i.e., $A \not\subseteq C$.

As before, we reason that $A \not\subseteq C$ iff not every $x \in A$ is also $\in C$, i.e., some $x \in A$ is $\notin C$. Don't worry, with practice you won't have to think hard anymore to unpack negations like this.

In other words, there is an x such that $x \in A$ and $x \notin C$.

Now we can use this to get to our contradiction. Of course, we'll have to use the other two assumptions to do it.

Since $A \subseteq B$, $x \in B$. Since $B \subseteq C$, $x \in C$. But this contradicts $x \notin C$. \square

Proposition 1.13. *If $A \cup B = A \cap B$ then $A = B$.*

Proof. Suppose $A \cup B = A \cap B$. We want to show that $A = B$.

The beginning is now routine:

Assume, by way of contradiction, that $A \neq B$.

Our assumption for the proof by contradiction is that $A \neq B$. Since $A = B$ iff $A \subseteq B$ and $B \subseteq A$, we get that $A \neq B$ iff $A \not\subseteq B$ or $B \not\subseteq A$. (Note how important it is to be careful when manipulating negations!) To prove a contradiction from this disjunction, we use a proof by cases and show that in each case, a contradiction follows.

$A \neq B$ iff $A \not\subseteq B$ or $B \not\subseteq A$. We distinguish cases.

In the first case, we assume $A \not\subseteq B$, i.e., for some x , $x \in A$ but $x \notin B$. $A \cap B$ is defined as those elements that A and B have in common, so if something isn't in one of them, it's not in the intersection. $A \cup B$ is A together with B , so anything in either is also in the union. This tells us that $x \in A \cup B$ but $x \notin A \cap B$, and hence that $A \cap B \neq A \cup B$.

Case 1: $A \not\subseteq B$. Then for some x , $x \in A$ but $x \notin B$. Since $x \notin B$, then $x \notin A \cap B$. Since $x \in A$, $x \in A \cup B$. So, $A \cap B \neq A \cup B$, contradicting the assumption that $A \cap B = A \cup B$.

Case 2: $B \not\subseteq A$. Then for some y , $y \in B$ but $y \notin A$. As before, we have $y \in A \cup B$ but $y \notin A \cap B$, and so $A \cap B \neq A \cup B$, again contradicting $A \cap B = A \cup B$. \square

1.8 Reading Proofs

math:prf:rea:
sec

Proofs you find in textbooks and articles very seldom give all the details we have so far included in our examples. Authors often do not draw attention to when they distinguish cases, when they give an indirect proof, or don't mention that they use a definition. So when you read a proof in a textbook, you will often have to fill in those details for yourself in order to understand the proof. Doing this is also good practice to get the hang of the various moves you have to make in a proof. Let's look at an example.

Proposition 1.14 (Absorption). *For all sets A , B ,*

$$A \cap (A \cup B) = A$$

Proof. If $z \in A \cap (A \cup B)$, then $z \in A$, so $A \cap (A \cup B) \subseteq A$. Now suppose $z \in A$. Then also $z \in A \cup B$, and therefore also $z \in A \cap (A \cup B)$. \square

The preceding proof of the absorption law is very condensed. There is no mention of any definitions used, no “we have to prove that” before we prove it, etc. Let’s unpack it. The proposition proved is a general claim about any sets A and B , and when the proof mentions A or B , these are variables for arbitrary sets. The general claim the proof establishes is what’s required to prove identity of sets, i.e., that every **element** of the left side of the identity is **an element** of the right and vice versa.

“If $z \in A \cap (A \cup B)$, then $z \in A$, so $A \cap (A \cup B) \subseteq A$.”

This is the first half of the proof of the identity: it establishes that if an arbitrary z is **an element** of the left side, it is also **an element** of the right, i.e., $A \cap (A \cup B) \subseteq A$. Assume that $z \in A \cap (A \cup B)$. Since z is an **element** of the intersection of two sets iff it is an **element** of both sets, we can conclude that $z \in A$ and also $z \in A \cup B$. In particular, $z \in A$, which is what we wanted to show. Since that’s all that has to be done for the first half, we know that the rest of the proof must be a proof of the second half, i.e., a proof that $A \subseteq A \cap (A \cup B)$.

“Now suppose $z \in A$. Then also $z \in A \cup B$, and therefore also $z \in A \cap (A \cup B)$.”

We start by assuming that $z \in A$, since we are showing that, for any z , if $z \in A$ then $z \in A \cap (A \cup B)$. To show that $z \in A \cap (A \cup B)$, we have to show (by definition of “ \cap ”) that (i) $z \in A$ and also (ii) $z \in A \cup B$. Here (i) is just our assumption, so there is nothing further to prove, and that’s why the proof does not mention it again. For (ii), recall that z is **an element** of a union of sets iff it is an **element** of at least one of those sets. Since $z \in A$, and $A \cup B$ is the union of A and B , this is the case here. So $z \in A \cup B$. We’ve shown both (i) $z \in A$ and (ii) $z \in A \cup B$, hence, by definition of “ \cap ,” $z \in A \cap (A \cup B)$. The proof doesn’t mention those definitions; it’s assumed the reader has already internalized them. If you haven’t, you’ll have to go back and remind yourself what they are. Then you’ll also have to recognize why it follows from $z \in A$ that $z \in A \cup B$, and from $z \in A$ and $z \in A \cup B$ that $z \in A \cap (A \cup B)$.

Here’s another version of the proof above, with everything made explicit:

Proof. [By definition of $=$ for sets, $A \cap (A \cup B) = A$ we have to show (a) $A \cap (A \cup B) \subseteq A$ and (b) $A \cap (A \cup B) \supseteq A$. (a): By definition of \subseteq , we have to show that if $z \in A \cap (A \cup B)$, then $z \in A$.] If $z \in A \cap (A \cup B)$, then $z \in A$ [since by definition of \cap , $z \in A \cap (A \cup B)$ iff $z \in A$ and $z \in A \cup B$], so $A \cap (A \cup B) \subseteq A$. [(b): By definition of \subseteq , we have to show that if $z \in A$, then $z \in A \cap (A \cup B)$.] Now suppose [(1)] $z \in A$. Then also [(2)] $z \in A \cup B$ [since by (1) $z \in A$ or $z \in B$, which by definition of \cup means $z \in A \cup B$], and therefore also $z \in A \cap (A \cup B)$ [since the definition of \cap requires that $z \in A$, i.e., (1), and $z \in A \cup B$, i.e., (2)]. \square

Problem 1.3. Expand the following proof of $A \cup (A \cap B) = A$, where you mention all the inference patterns used, why each step follows from assumptions or claims established before it, and where we have to appeal to which definitions.

Proof. If $z \in A \cup (A \cap B)$ then $z \in A$ or $z \in A \cap B$. If $z \in A \cap B$, $z \in A$. Any $z \in A$ is also $\in A \cup (A \cap B)$. \square

1.9 I Can't Do It!

math:prf:cnt:
sec

We all get to a point where we feel like giving up. But you *can* do it. Your instructor and teaching assistant, as well as your fellow students, can help. Ask them for help! Here are a few tips to help you avoid a crisis, and what to do if you feel like giving up.

To make sure you can solve problems successfully, do the following:

1. *Start as far in advance as possible.* We get busy throughout the semester and many of us struggle with procrastination, one of the best things you can do is to start your homework assignments early. That way, if you're stuck, you have time to look for a solution (that isn't crying).
2. *Talk to your classmates.* You are not alone. Others in the class may also struggle—but they may struggle with different things. Talking it out with your peers can give you a different perspective on the problem that might lead to a breakthrough. Of course, don't just copy their solution: ask them for a hint, or explain where you get stuck and ask them for the next step. And when you do get it, reciprocate. Helping someone else along, and explaining things will help you understand better, too.
3. *Ask for help.* You have many resources available to you—your instructor and teaching assistant are there for you and *want* you to succeed. They should be able to help you work out a problem and identify where in the process you're struggling.
4. *Take a break.* If you're stuck, it *might* be because you've been staring at the problem for too long. Take a short break, have a cup of tea, or work on a different problem for a while, then return to the problem with a fresh mind. Sleep on it.

Notice how these strategies require that you've started to work on the proof well in advance? If you've started the proof at 2am the day before it's due, these might not be so helpful.

This might sound like doom and gloom, but solving a proof is a challenge that pays off in the end. Some people do this as a career—so there must be something to enjoy about it. Like basically everything, solving problems and doing proofs is something that requires practice. You might see classmates who find this easy: they've probably just had lots of practice already. Try not to give in too easily.

If you do run out of time (or patience) on a particular problem: that's ok. It doesn't mean you're stupid or that you will never get it. Find out (from your instructor or another student) how it is done, and identify where you went wrong or got stuck, so you can avoid doing that the next time you encounter a similar issue. Then try to do it without looking at the solution. And next time, start (and ask for help) earlier.

1.10 Other Resources

There are many books on how to do proofs in mathematics which may be useful. Check out *How to Read and do Proofs: An Introduction to Mathematical Thought Processes* (Solow, 2013) and *How to Prove It: A Structured Approach* (Velleman, 2019) in particular. The *Book of Proof* (Hammack, 2013) and *Mathematical Reasoning* (Sandstrum, 2019) are books on proof that are freely available online. Philosophers might find *More Precisely: The Math you need to do Philosophy* (Steinhart, 2018) to be a good primer on mathematical reasoning.

There are also various shorter guides to proofs available on the internet; e.g., “Introduction to Mathematical Arguments” (Hutchings, 2003) and “How to write proofs” (Cheng, 2004).

Motivational Videos

Feel like you have no motivation to do your homework? Feeling down? These videos might help!

- https://www.youtube.com/watch?v=ZXsQAXx_ao0
- <https://www.youtube.com/watch?v=BQ4yd2W50No>
- <https://www.youtube.com/watch?v=StTqXEQ21-Y>

Chapter 2

Induction

2.1 Introduction

mth:ind:int:
sec Induction is an important proof technique which is used, in different forms, in almost all areas of logic, theoretical computer science, and mathematics. It is needed to prove many of the results in logic.

Induction is often contrasted with deduction, and characterized as the inference from the particular to the general. For instance, if we observe many green emeralds, and nothing that we would call an emerald that's not green, we might conclude that all emeralds are green. This is an inductive inference, in that it proceeds from many particular cases (this emerald is green, that emerald is green, etc.) to a general claim (all emeralds are green). *Mathematical induction* is also an inference that concludes a general claim, but it is of a very different kind than this “simple induction.”

Very roughly, an inductive proof in mathematics concludes that all mathematical objects of a certain sort have a certain property. In the simplest case, the mathematical objects an inductive proof is concerned with are natural numbers. In that case an inductive proof is used to establish that all natural numbers have some property, and it does this by showing that

1. 0 has the property, and
2. whenever a number k has the property, so does $k + 1$.

Induction on natural numbers can then also often be used to prove general claims about mathematical objects that can be assigned numbers. For instance, finite sets each have a finite number n of elements, and if we can use induction to show that every number n has the property “all finite sets of size n are . . .” then we will have shown something about all finite sets.

Induction can also be generalized to mathematical objects that are *inductively defined*. For instance, expressions of a formal language such as those of first-order logic are defined inductively. *Structural induction* is a way to prove results about all such expressions. Structural induction, in particular, is very useful—and widely used—in logic.

2.2 Induction on \mathbb{N}

In its simplest form, induction is a technique used to prove results for all natural numbers. It uses the fact that by starting from 0 and repeatedly adding 1 we eventually reach every natural number. So to prove that something is true for every number, we can (1) establish that it is true for 0 and (2) show that whenever it is true for a number n , it is also true for the next number $n + 1$. If we abbreviate “number n has property P ” by $P(n)$ (and “number k has property P ” by $P(k)$, etc.), then a proof by induction that $P(n)$ for all $n \in \mathbb{N}$ consists of:

1. a proof of $P(0)$, and
2. a proof that, for any k , if $P(k)$ then $P(k + 1)$.

To make this crystal clear, suppose we have both (1) and (2). Then (1) tells us that $P(0)$ is true. If we also have (2), we know in particular that if $P(0)$ then $P(0 + 1)$, i.e., $P(1)$. This follows from the general statement “for any k , if $P(k)$ then $P(k + 1)$ ” by putting 0 for k . So by modus ponens, we have that $P(1)$. From (2) again, now taking 1 for n , we have: if $P(1)$ then $P(2)$. Since we’ve just established $P(1)$, by modus ponens, we have $P(2)$. And so on. For any number n , after doing this n times, we eventually arrive at $P(n)$. So (1) and (2) together establish $P(n)$ for any $n \in \mathbb{N}$.

Let’s look at an example. Suppose we want to find out how many different sums we can throw with n dice. Although it might seem silly, let’s start with 0 dice. If you have no dice there’s only one possible sum you can “throw”: no dots at all, which sums to 0. So the number of different possible throws is 1. If you have only one die, i.e., $n = 1$, there are six possible values, 1 through 6. With two dice, we can throw any sum from 2 through 12, that’s 11 possibilities. With three dice, we can throw any number from 3 to 18, i.e., 16 different possibilities. 1, 6, 11, 16: looks like a pattern: maybe the answer is $5n + 1$? Of course, $5n + 1$ is the maximum possible, because there are only $5n + 1$ numbers between n , the lowest value you can throw with n dice (all 1’s) and $6n$, the highest you can throw (all 6’s).

Theorem 2.1. *With n dice one can throw all $5n + 1$ possible values between n and $6n$.*

Proof. Let $P(n)$ be the claim: “It is possible to throw any number between n and $6n$ using n dice.” To use induction, we prove:

1. The *induction basis* $P(1)$, i.e., with just one die, you can throw any number between 1 and 6.
2. The *induction step*, for all k , if $P(k)$ then $P(k + 1)$.

(1) Is proved by inspecting a 6-sided die. It has all 6 sides, and every number between 1 and 6 shows up one on of the sides. So it is possible to throw any number between 1 and 6 using a single die.

To prove (2), we assume the antecedent of the conditional, i.e., $P(k)$. This assumption is called the *inductive hypothesis*. We use it to prove $P(k+1)$. The hard part is to find a way of thinking about the possible values of a throw of $k+1$ dice in terms of the possible values of throws of k dice plus of throws of the extra $k+1$ -st die—this is what we have to do, though, if we want to use the inductive hypothesis.

The inductive hypothesis says we can get any number between k and $6k$ using k dice. If we throw a 1 with our $(k+1)$ -st die, this adds 1 to the total. So we can throw any value between $k+1$ and $6k+1$ by throwing k dice and then rolling a 1 with the $(k+1)$ -st die. What's left? The values $6k+2$ through $6k+6$. We can get these by rolling k 6s and then a number between 2 and 6 with our $(k+1)$ -st die. Together, this means that with $k+1$ dice we can throw any of the numbers between $k+1$ and $6(k+1)$, i.e., we've proved $P(k+1)$ using the assumption $P(k)$, the inductive hypothesis. \square

Very often we use induction when we want to prove something about a series of objects (numbers, sets, etc.) that is itself defined “inductively,” i.e., by defining the $(n+1)$ -st object in terms of the n -th. For instance, we can define the sum s_n of the natural numbers up to n by

$$\begin{aligned}s_0 &= 0 \\ s_{n+1} &= s_n + (n+1)\end{aligned}$$

This definition gives:

$$\begin{aligned}s_0 &= 0, \\ s_1 &= s_0 + 1 &&= 1, \\ s_2 &= s_1 + 2 &&= 1 + 2 = 3 \\ s_3 &= s_2 + 3 &&= 1 + 2 + 3 = 6, \text{ etc.}\end{aligned}$$

Now we can prove, by induction, that $s_n = n(n+1)/2$.

Proposition 2.2. $s_n = n(n+1)/2$.

Proof. We have to prove (1) that $s_0 = 0 \cdot (0+1)/2$ and (2) if $s_k = k(k+1)/2$ then $s_{k+1} = (k+1)(k+2)/2$. (1) is obvious. To prove (2), we assume the inductive hypothesis: $s_k = k(k+1)/2$. Using it, we have to show that $s_{k+1} = (k+1)(k+2)/2$.

What is s_{k+1} ? By the definition, $s_{k+1} = s_k + (k+1)$. By inductive hypothesis, $s_k = k(k+1)/2$. We can substitute this into the previous equation,

and then just need a bit of arithmetic of fractions:

$$\begin{aligned}
 s_{k+1} &= \frac{k(k+1)}{2} + (k+1) = \\
 &= \frac{k(k+1)}{2} + \frac{2(k+1)}{2} = \\
 &= \frac{k(k+1) + 2(k+1)}{2} = \\
 &= \frac{(k+2)(k+1)}{2}. \quad \square
 \end{aligned}$$

The important lesson here is that if you’re proving something about some inductively defined sequence a_n , induction is the obvious way to go. And even if it isn’t (as in the case of the possibilities of dice throws), you can use induction if you can somehow relate the case for $k+1$ to the case for k .

2.3 Strong Induction

In the principle of induction discussed above, we prove $P(0)$ and also if $P(k)$, then $P(k+1)$. In the second part, we assume that $P(k)$ is true and use this assumption to prove $P(k+1)$. Equivalently, of course, we could assume $P(k-1)$ and use it to prove $P(k)$ —the important part is that we be able to carry out the inference from any number to its successor; that we can prove the claim in question for any number under the assumption it holds for its predecessor. meth:ind:str:
sec

There is a variant of the principle of induction in which we don’t just assume that the claim holds for the predecessor $k-1$ of k , but for all numbers smaller than k , and use this assumption to establish the claim for k . This also gives us the claim $P(n)$ for all $n \in \mathbb{N}$. For once we have established $P(0)$, we have thereby established that P holds for all numbers less than 1. And if we know that if $P(l)$ for all $l < k$, then $P(k)$, we know this in particular for $k=1$. So we can conclude $P(1)$. With this we have proved $P(0)$ and $P(1)$, i.e., $P(l)$ for all $l < 2$, and since we have also the conditional, if $P(l)$ for all $l < 2$, then $P(2)$, we can conclude $P(2)$, and so on.

In fact, if we can establish the general conditional “for all k , if $P(l)$ for all $l < k$, then $P(k)$,” we do not have to establish $P(0)$ anymore, since it follows from it. For remember that a general claim like “for all $l < k$, $P(l)$ ” is true if there are no $l < k$. This is a case of vacuous quantification: “all As are Bs ” is true if there are no As , $\forall x (\varphi(x) \rightarrow \psi(x))$ is true if no x satisfies $\varphi(x)$. In this case, the formalized version would be “ $\forall l (l < k \rightarrow P(l))$ ”—and that is true if there are no $l < k$. And if $k=0$ that’s exactly the case: no $l < 0$, hence “for all $l < 0$, $P(0)$ ” is true, whatever P is. A proof of “if $P(l)$ for all $l < k$, then $P(k)$ ” thus automatically establishes $P(0)$.

This variant is useful if establishing the claim for k can’t be made to just rely on the claim for $k-1$ but may require the assumption that it is true for one or more $l < k$.

2.4 Inductive Definitions

math.ind:df:
sec

In logic we very often define kinds of objects *inductively*, i.e., by specifying rules for what counts as an object of the kind to be defined which explain how to get new objects of that kind from old objects of that kind. For instance, we often define special kinds of sequences of symbols, such as the terms and **formulas** of a language, by induction. For a simple example, consider strings of consisting of letters a, b, c, d, the symbol \circ , and brackets [and], such as “[c \circ d]”, “[a[] \circ]”, “a” or “[a \circ b] \circ d]”. You probably feel that there’s something “wrong” with the first two strings: the brackets don’t “balance” at all in the first, and you might feel that the “ \circ ” should “connect” expressions that themselves make sense. The third and fourth string look better: for every “[” there’s a closing “]” (if there are any at all), and for any \circ we can find “nice” expressions on either side, surrounded by a pair of parentheses.

We would like to precisely specify what counts as a “nice term.” First of all, every letter by itself is nice. Anything that’s not just a letter by itself should be of the form “[$t \circ s$]” where s and t are themselves nice. Conversely, if t and s are nice, then we can form a new nice term by putting a \circ between them and surround them by a pair of brackets. We might use these operations to *define* the set of nice terms. This is an *inductive definition*.

Definition 2.3 (Nice terms). The set of *nice terms* is inductively defined as follows:

1. Any letter a, b, c, d is a nice term.
2. If s_1 and s_2 are nice terms, then so is [$s_1 \circ s_2$].
3. Nothing else is a nice term.

This definition tells us that something counts as a nice term iff it can be constructed according to the two conditions (1) and (2) in some finite number of steps. In the first step, we construct all nice terms just consisting of letters by themselves, i.e.,

a, b, c, d

In the second step, we apply (2) to the terms we’ve constructed. We’ll get

[a \circ a], [a \circ b], [b \circ a], . . . , [d \circ d]

for all combinations of two letters. In the third step, we apply (2) again, to any two nice terms we’ve constructed so far. We get new nice term such as [a \circ [a \circ a]]—where t is a from step 1 and s is [a \circ a] from step 2—and [[b \circ c] \circ [d \circ b]] constructed out of the two terms [b \circ c] and [d \circ b] from step 2. And so on. Clause (3) rules out that anything not constructed in this way sneaks into the set of nice terms.

Note that we have not yet proved that every sequence of symbols that “feels” nice is nice according to this definition. However, it should be clear that

everything we can construct does in fact “feel nice”: brackets are balanced, and \circ connects parts that are themselves nice.

The key feature of inductive definitions is that if you want to prove something about all nice terms, the definition tells you which cases you must consider. For instance, if you are told that t is a nice term, the inductive definition tells you what t can look like: t can be a letter, or it can be $[s_1 \circ s_2]$ for some pair of nice terms s_1 and s_2 . Because of clause (3), those are the only possibilities.

When proving claims about all of an inductively defined set, the strong form of induction becomes particularly important. For instance, suppose we want to prove that for every nice term of length n , the number of $[$ in it is $< n/2$. This can be seen as a claim about all n : for every n , the number of $[$ in any nice term of length n is $< n/2$.

Proposition 2.4. *For any n , the number of $[$ in a nice term of length n is $< n/2$.*

Proof. To prove this result by (strong) induction, we have to show that the following conditional claim is true:

If for every $l < k$, any nice term of length l has $< l/2$ $[$'s, then any nice term of length k has $< k/2$ $[$'s.

To show this conditional, assume that its antecedent is true, i.e., assume that for any $l < k$, nice terms of length l contain $< l/2$ $[$'s. We call this assumption the inductive hypothesis. We want to show the same is true for nice terms of length k .

So suppose t is a nice term of length k . Because nice terms are inductively defined, we have two cases: (1) t is a letter by itself, or (2) t is $[s_1 \circ s_2]$ for some nice terms s_1 and s_2 .

1. t is a letter. Then $k = 1$, and the number of $[$ in t is 0. Since $0 < 1/2$, the claim holds.
2. t is $[s_1 \circ s_2]$ for some nice terms s_1 and s_2 . Let's let l_1 be the length of s_1 and l_2 be the length of s_2 . Then the length k of t is $l_1 + l_2 + 3$ (the lengths of s_1 and s_2 plus three symbols $[$, \circ , $]$). Since $l_1 + l_2 + 3$ is always greater than l_1 , $l_1 < k$. Similarly, $l_2 < k$. That means that the induction hypothesis applies to the terms s_1 and s_2 : the number m_1 of $[$ in s_1 is $< l_1/2$, and the number m_2 of $[$ in s_2 is $< l_2/2$.

The number of $[$ in t is the number of $[$ in s_1 , plus the number of $[$ in s_2 , plus 1, i.e., it is $m_1 + m_2 + 1$. Since $m_1 < l_1/2$ and $m_2 < l_2/2$ we have:

$$m_1 + m_2 + 1 < \frac{l_1}{2} + \frac{l_2}{2} + 1 = \frac{l_1 + l_2 + 2}{2} < \frac{l_1 + l_2 + 3}{2} = k/2.$$

In each case, we've shown that the number of $[$ in t is $< k/2$ (on the basis of the inductive hypothesis). By strong induction, the proposition follows. \square

Problem 2.1. Define the set of supernice terms by

1. Any letter a, b, c, d is a supernice term.
2. If s is a supernice term, then so is $[s]$.
3. If s_1 and s_2 are supernice terms, then so is $[s_1 \circ s_2]$.
4. Nothing else is a supernice term.

Show that the number of $[$ in a supernice term t of length n is $\leq n/2 + 1$.

2.5 Structural Induction

math:ind:sti:
sec So far we have used induction to establish results about all natural numbers. But a corresponding principle can be used directly to prove results about all **elements** of an inductively defined set. This is often called *structural* induction, because it depends on the structure of the inductively defined objects.

Generally, an inductive definition is given by (a) a list of “initial” **elements** of the set and (b) a list of operations which produce new **elements** of the set from old ones. In the case of nice terms, for instance, the initial objects are the letters. We only have one operation: the operations are

$$o(s_1, s_2) = [s_1 \circ s_2]$$

You can even think of the natural numbers \mathbb{N} themselves as being given by an inductive definition: the initial object is 0, and the operation is the successor function $x + 1$.

In order to prove something about all elements of an inductively defined set, i.e., that every **element** of the set has a property P , we must:

1. Prove that the initial objects have P
2. Prove that for each operation o , if the arguments have P , so does the result.

For instance, in order to prove something about all nice terms, we would prove that it is true about all letters, and that it is true about $[s_1 \circ s_2]$ provided it is true of s_1 and s_2 individually.

Proposition 2.5. *The number of $[$ equals the number of $]$ in any nice term t .*

Proof. We use structural induction. Nice terms are inductively defined, with letters as initial objects and the operation o for constructing new nice terms out of old ones.

1. The claim is true for every letter, since the number of $[$ in a letter by itself is 0 and the number of $]$ in it is also 0.

2. Suppose the number of $[$ in s_1 equals the number of $]$, and the same is true for s_2 . The number of $[$ in $o(s_1, s_2)$, i.e., in $[s_1 \circ s_2]$, is the sum of the number of $[$ in s_1 and s_2 plus one. The number of $]$ in $o(s_1, s_2)$ is the sum of the number of $]$ in s_1 and s_2 plus one. Thus, the number of $[$ in $o(s_1, s_2)$ equals the number of $]$ in $o(s_1, s_2)$. \square

Problem 2.2. Prove by structural induction that no nice term starts with $]$.

Let's give another proof by structural induction: a proper initial segment of a string t of symbols is any string s that agrees with t symbol by symbol, read from the left, but t is longer. So, e.g., $[a \circ$ is a proper initial segment of $[a \circ b]$, but neither are $[b \circ$ (they disagree at the second symbol) nor $[a \circ b]$ (they are the same length).

Proposition 2.6. *Every proper initial segment of a nice term t has more $[$'s than $]$'s.* mth:ind:sti:
prop:initial

Proof. By induction on t :

1. t is a letter by itself: Then t has no proper initial segments.
2. $t = [s_1 \circ s_2]$ for some nice terms s_1 and s_2 . If r is a proper initial segment of t , there are a number of possibilities:
 - a) r is just $[$: Then r has one more $[$ than it does $]$.
 - b) r is $[r_1$ where r_1 is a proper initial segment of s_1 : Since s_1 is a nice term, by induction hypothesis, r_1 has more $[$ than $]$ and the same is true for $[r_1$.
 - c) r is $[s_1$ or $[s_1 \circ$: By the previous result, the number of $[$ and $]$ in s_1 are equal; so the number of $[$ in $[s_1$ or $[s_1 \circ$ is one more than the number of $]$.
 - d) r is $[s_1 \circ r_2$ where r_2 is a proper initial segment of s_2 : By induction hypothesis, r_2 contains more $[$ than $]$. By the previous result, the number of $[$ and of $]$ in s_1 are equal. So the number of $[$ in $[s_1 \circ r_2$ is greater than the number of $]$.
 - e) r is $[s_1 \circ s_2$: By the previous result, the number of $[$ and $]$ in s_1 are equal, and the same for s_2 . So there is one more $[$ in $[s_1 \circ s_2$ than there are $]$. \square

2.6 Relations and Functions

When we have defined a set of objects (such as the natural numbers or the nice terms) inductively, we can also define *relations on* these objects by induction. For instance, consider the following idea: a nice term t_1 is a subterm of a nice term t_2 if it occurs as a part of it. Let's use a symbol for it: $t_1 \sqsubseteq t_2$. Every mth:ind:rel:
sec

nice term is a subterm of itself, of course: $t \sqsubseteq t$. We can give an inductive definition of this relation as follows:

Definition 2.7. The relation of a nice term t_1 being a subterm of t_2 , $t_1 \sqsubseteq t_2$, is defined by induction on t_2 as follows:

1. If t_2 is a letter, then $t_1 \sqsubseteq t_2$ iff $t_1 = t_2$.
2. If t_2 is $[s_1 \circ s_2]$, then $t_1 \sqsubseteq t_2$ iff $t_1 = t_2$, $t_1 \sqsubseteq s_1$, or $t_1 \sqsubseteq s_2$.

This definition, for instance, will tell us that $a \sqsubseteq [b \circ a]$. For (2) says that $a \sqsubseteq [b \circ a]$ iff $a = [b \circ a]$, or $a \sqsubseteq b$, or $a \sqsubseteq a$. The first two are false: a clearly isn't identical to $[b \circ a]$, and by (1), $a \sqsubseteq b$ iff $a = b$, which is also false. However, also by (1), $a \sqsubseteq a$ iff $a = a$, which is true.

It's important to note that the success of this definition depends on a fact that we haven't proved yet: every nice term t is either a letter by itself, or there are *uniquely determined* nice terms s_1 and s_2 such that $t = [s_1 \circ s_2]$. "Uniquely determined" here means that if $t = [s_1 \circ s_2]$ it isn't *also* $= [r_1 \circ r_2]$ with $s_1 \neq r_1$ or $s_2 \neq r_2$. If this were the case, then clause (2) may come in conflict with itself: reading t_2 as $[s_1 \circ s_2]$ we might get $t_1 \sqsubseteq t_2$, but if we read t_2 as $[r_1 \circ r_2]$ we might get not $t_1 \sqsubseteq t_2$. Before we prove that this can't happen, let's look at an example where it *can* happen.

Definition 2.8. Define *bracketless terms* inductively by

1. Every letter is a bracketless term.
2. If s_1 and s_2 are bracketless terms, then $s_1 \circ s_2$ is a bracketless term.
3. Nothing else is a bracketless term.

Bracketless terms are, e.g., a , $b \circ d$, $b \circ a \circ b$. Now if we defined "subterm" for bracketless terms the way we did above, the second clause would read

If $t_2 = s_1 \circ s_2$, then $t_1 \sqsubseteq t_2$ iff $t_1 = t_2$, $t_1 \sqsubseteq s_1$, or $t_1 \sqsubseteq s_2$.

Now $b \circ a \circ b$ is of the form $s_1 \circ s_2$ with

$$s_1 = b \text{ and } s_2 = a \circ b.$$

It is also of the form $r_1 \circ r_2$ with

$$r_1 = b \circ a \text{ and } r_2 = b.$$

Now is $a \circ b$ a subterm of $b \circ a \circ b$? The answer is yes if we go by the first reading, and no if we go by the second.

The property that the way a nice term is built up from other nice terms is unique is called *unique readability*. Since inductive definitions of relations for such inductively defined objects are important, we have to prove that it holds.

Proposition 2.9. *Suppose t is a nice term. Then either t is a letter by itself, or there are uniquely determined nice terms s_1, s_2 such that $t = [s_1 \circ s_2]$.*

Proof. If t is a letter by itself, the condition is satisfied. So assume t isn't a letter by itself. We can tell from the inductive definition that then t must be of the form $[s_1 \circ s_2]$ for some nice terms s_1 and s_2 . It remains to show that these are uniquely determined, i.e., if $t = [r_1 \circ r_2]$, then $s_1 = r_1$ and $s_2 = r_2$.

So suppose $t = [s_1 \circ s_2]$ and also $t = [r_1 \circ r_2]$ for nice terms s_1, s_2, r_1, r_2 . We have to show that $s_1 = r_1$ and $s_2 = r_2$. First, s_1 and r_1 must be identical, for otherwise one is a proper initial segment of the other. But by [Proposition 2.6](#), that is impossible if s_1 and r_1 are both nice terms. But if $s_1 = r_1$, then clearly also $s_2 = r_2$. \square

We can also define functions inductively: e.g., we can define the function f that maps any nice term to the maximum depth of nested $[\dots]$ in it as follows:

Definition 2.10. The *depth* of a nice term, $f(t)$, is defined inductively as follows: [mth:ind:rel:](#)
[defn:depth](#)

$$f(t) = \begin{cases} 0 & \text{if } t \text{ is a letter} \\ \max(f(s_1), f(s_2)) + 1 & \text{if } t = [s_1 \circ s_2]. \end{cases}$$

For instance

$$\begin{aligned} f([a \circ b]) &= \max(f(a), f(b)) + 1 = \\ &= \max(0, 0) + 1 = 1, \text{ and} \\ f([[a \circ b] \circ c]) &= \max(f([a \circ b]), f(c)) + 1 = \\ &= \max(1, 0) + 1 = 2. \end{aligned}$$

Here, of course, we assume that s_1 and s_2 are nice terms, and make use of the fact that every nice term is either a letter or of the form $[s_1 \circ s_2]$. It is again important that it can be of this form in only one way. To see why, consider again the bracketless terms we defined earlier. The corresponding “definition” would be:

$$g(t) = \begin{cases} 0 & \text{if } t \text{ is a letter} \\ \max(g(s_1), g(s_2)) + 1 & \text{if } t = s_1 \circ s_2. \end{cases}$$

Now consider the bracketless term $a \circ b \circ c \circ d$. It can be read in more than one way, e.g., as $s_1 \circ s_2$ with

$$s_1 = a \text{ and } s_2 = b \circ c \circ d,$$

or as $r_1 \circ r_2$ with

$$r_1 = a \circ b \text{ and } r_2 = c \circ d.$$

Calculating g according to the first way of reading it would give

$$\begin{aligned}g(s_1 \circ s_2) &= \max(g(a), g(b \circ c \circ d)) + 1 = \\ &= \max(0, 2) + 1 = 3\end{aligned}$$

while according to the other reading we get

$$\begin{aligned}g(r_1 \circ r_2) &= \max(g(a \circ b), g(c \circ d)) + 1 = \\ &= \max(1, 1) + 1 = 2\end{aligned}$$

But a function must always yield a unique value; so our “definition” of g doesn’t define a function at all.

Problem 2.3. Give an inductive definition of the function l , where $l(t)$ is the number of symbols in the nice term t .

Problem 2.4. Prove by structural induction on nice terms t that $f(t) < l(t)$ (where $l(t)$ is the number of symbols in t and $f(t)$ is the depth of t as defined in [Definition 2.10](#)).

Photo Credits

Bibliography

- Cheng, Eugenia. 2004. How to write proofs: A quick guide. URL <http://http://eugeniacheng.com/wp-content/uploads/2017/02/cheng-proofguide.pdf>.
- Hammack, Richard. 2013. *Book of Proof*. Richmond, VA: Virginia Commonwealth University. URL <http://www.people.vcu.edu/~rhammack/BookOfProof/BookOfProof.pdf>.
- Hutchings, Michael. 2003. Introduction to mathematical arguments. URL <https://math.berkeley.edu/~hutching/teach/proofs.pdf>.
- Sandstrum, Ted. 2019. *Mathematical Reasoning: Writing and Proof*. Allendale, MI: Grand Valley State University. URL <https://scholarworks.gvsu.edu/books/7/>.
- Solow, Daniel. 2013. *How to Read and Do Proofs*. Hoboken, NJ: Wiley.
- Steinhart, Eric. 2018. *More Precisely: The Math You Need to Do Philosophy*. Peterborough, ON: Broadview, 2nd ed.
- Velleman, Daniel J. 2019. *How to Prove It: A Structured Approach*. Cambridge: Cambridge University Press, 3rd ed.