Lecture Notes: Week 6

Finding Proofs

This week we are going to shift gears a little bit. Instead of introducing more theory, we are going to prove a few statements using the theory we have already built. Our goal, however, is not to jump from the problem statement directly to a working proof (similar to how the "Definition, Theorem, proof" style works). Instead, we will spend a bit of time thinking about how to find and construct a proof.

Generally speaking, it works like this (as many of you have discovered in the past few weeks):

1. Take a moment to digest the statement.
2. Ask if it feels intuitively true.
3. If so, use definitions & theorems to pull apart the statement.
4. If not, prod the statement. What seems odd or important? Use definitions and theorems to try to determine why.

Problem 1

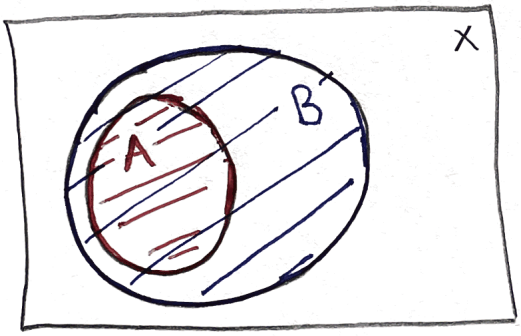
Let A and B be subsets of a set X.

$A \subset B$ if and only if $B^c \subset A^c$, where "^{c}" indicates a complement relative to X.

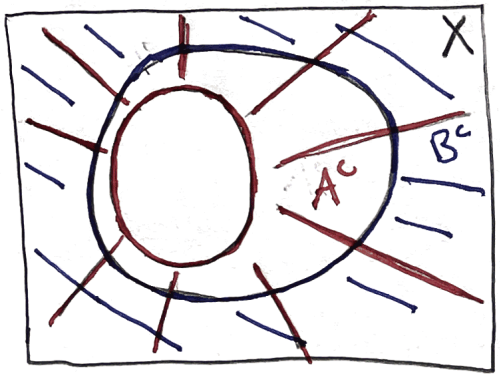
Sketch / Finding a proof:

First, does the statement seem true? Perhaps not, but if we try to draw this... it might seem clearer.

Make $A \subset B$:



and notice, if we take complements:



looks like $B^c \subset A^c$.

Perhaps this feels more intuitive now, so we should feel confident trying to prove it.

To find the proof, we begin prodding the statement. We should grab the objects and statements, and use our definitions and theorems to unravel them.

For instance, ① $A \subset B$ means if $x \in A$, then $x \in B$. DEF

Similarly, ② $B^c \subset A^c$ means if $x \in B^c$, then $x \in A^c$. DEF

We also know: ③ A^c is the set of $x \in X$ such that $x \notin A$ DEF

and similarly, ④ B^c is the set of $x \in X$ such that $x \notin B$. DEF

Notice, we can use ④ to restate ②: take $x \in X$, and

⑤ $B^c \subset A^c$ means if $x \notin B$, then $x \notin A$.

Now, statement ⑤ and statement ① look very close...

in fact...

"if $x \notin B$, then $x \notin A$"

is the contrapositive of

"if $x \in A$, then $x \in B$",

in other words, these are logically equivalent. Now, if you think about it for a moment, you will realize we have all of the pieces we need:

$$\begin{aligned} A \subset B &\iff \text{if } x \in A, \text{ then } x \in B \\ &\iff \text{if } x \notin B, \text{ then } x \notin A \\ &\iff \text{if } x \in B^c, \text{ then } x \in A^c \\ &\iff B^c \subset A^c \end{aligned}$$

However, we are not quite done. We have all of the elements we need to write a good proof, but we still need to write it. The process of writing is, well, perhaps not as fun as the process of discovery, or finding the proof. But, alas, to be a mathematician, we must carefully write our discovery.

proof: Assume $A \subset B$ and let $x \in X$. If $x \in A$, then $x \in B$, by definition of subset. Taking the contrapositive of this statement yields, "If $x \notin B$, then $x \notin A$." By definition of a complement relative to X, we see that this means if $x \in B^c$, then $x \in A^c$. By definition of subset, this means $B^c \subset A^c$, as desired. ▨

Notice that in our written proof, we only used the necessary information from our sketch, we did not draw any pictures, and each sentence was justified by machinery we have (definitions, theorems, etc.) More on writing later. For now, let's try another problem. A slightly more difficult problem.

Problem 2

Let A and B be sets. Show that if $\mathcal{S}(A \cup B) = \mathcal{S}(A) \cup \mathcal{S}(B)$, then $A \subset B$ or $B \subset A$.

($\mathcal{S}$: power set!)

Sketch / Finding a proof:

First, does the statement feel intuitively true? Perhaps not... maybe, though. For me, it does not - we have no theorems telling us about how unions interact with power sets.

In other words, "$\mathcal{S}(A \cup B) = \mathcal{S}(A) \cup \mathcal{S}(B)$" is not telling me much. We should focus on this set equality.

Somehow, we need to **deduce** something from it...

Let's try to pull it apart. First, let's use the definition of the power set

$\mathcal{P}(A \cup B)$ is the set of all subsets of $A \cup B$, which means if $S \in \mathcal{P}(A \cup B)$, then $S \subset A \cup B$. Similarly, we can pull apart $\mathcal{P}(A) \cup \mathcal{P}(B)$. This is a union of two power sets, so $S \in \mathcal{P}(A) \cup \mathcal{P}(B)$ means $S \in \mathcal{P}(A)$ or $S \subset \mathcal{P}(B)$, by definition of the union.

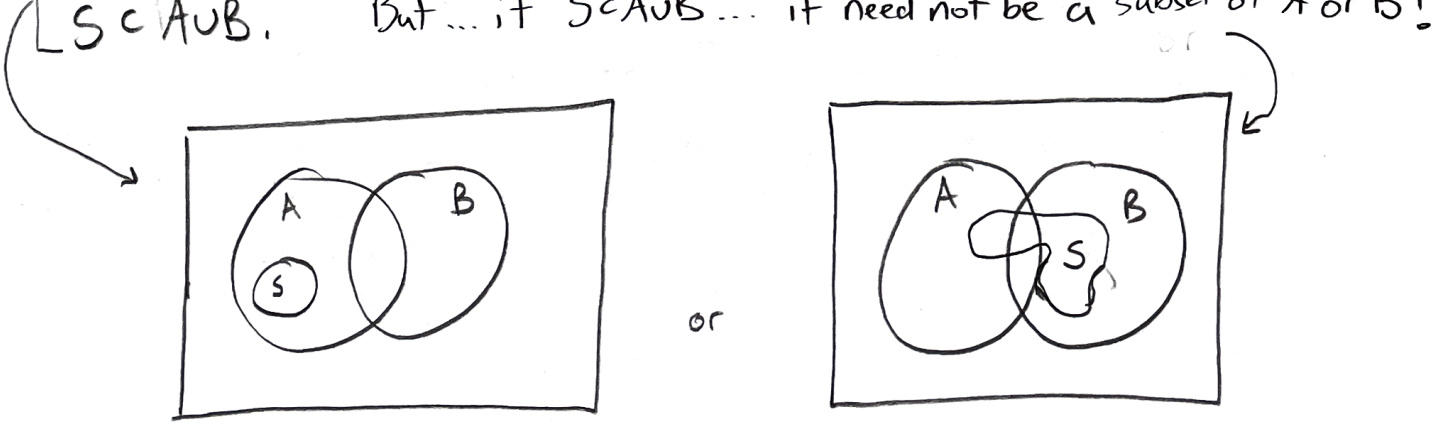
Continuing, that means $S \subset A$ or $S \subset B$.

Now, lets think about thiese to sets

① $S \in \mathcal{P}(A \cup B)$ means $S \subset A \cup B$

② $S \in \mathcal{P}(A) \cup \mathcal{P}(B)$ means $S \subset A$ or $S \subset B$.

If $S \subset A$, then $S \subset A \cup B$,.. similarly, if $S \subset B$, then $S \subset A \cup B$. But... if $S \subset A \cup B$... it need not be a subset of A or B!



Piecing this together, we can conclude this:

$$S \in \mathcal{P}(A) \cup \mathcal{P}(B) \iff S \subset A \text{ or } S \subset B$$
$$\Rightarrow S \subset A \cup B$$
$$\Leftrightarrow S \in \mathcal{P}(A \cup B).$$

In other words, using the definition of a subset, we see $\mathcal{P}(A) \cup \mathcal{P}(B) \subset \mathcal{P}(A \cup B)$ ... always!

We picked $S \in \mathcal{P}(A) \cup \mathcal{P}(B)$ arbitrarily (without conditions), so every element S in $\mathcal{P}(A) \cup \mathcal{P}(B)$ is also in $\mathcal{P}(A \cup B)$.

(*) [This is an important discovery! And the crazy part is ... we won't even need to mention it in our final proof of the statement!] (*)

Why is it important? Remember, we are prodding the statement $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$, since this is our assumption. Set equality is equivalent to $\mathcal{P}(A \cup B) \subset \mathcal{P}(A) \cup \mathcal{P}(B)$ and $\mathcal{P}(A) \cup \mathcal{P}(B) \subset \mathcal{P}(A \cup B)$, by THM 1 in the Week #4 notes.

What we just discovered is that one of these statements, namely $\mathcal{S}(A) \cup \mathcal{S}(B) \subset \mathcal{S}(A \cup B)$, always holds. If we take $\mathcal{S}(A) \cup \mathcal{S}(B) \subset \mathcal{S}(A \cup B)$ as a hypothesis, whatever the conclusion is must also always hold ... so its not going to restrict the conclusion.

In other words, it is this part of the equality which holds information that restricts properties of the sets A and B:

$$\mathcal{S}(A \cup B) \subset \mathcal{S}(A) \cup \mathcal{S}(B).$$

For this to hold, there must be something special about A and B which we can conclude.

Basically, our statement reduces to the following:

If $\mathcal{S}(A \cup B) = \mathcal{S}(A) \cup \mathcal{S}(B)$, then $A \subset B$ or $B \subset A$.

becomes

If $\mathcal{S}(A \cup B) \subset \mathcal{S}(A) \cup \mathcal{S}(B)$, then $A \subset B$ or $B \subset A$.

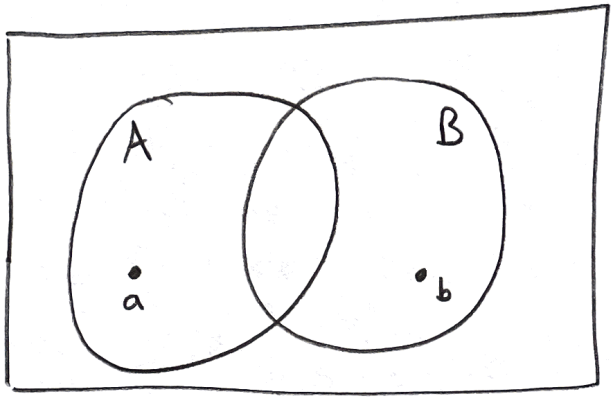
Now, one way forward would be to attempt to prove this directly. Namely, using $S \in \mathcal{S}(A \cup B) \Rightarrow S \in \mathcal{S}(A) \cup \mathcal{S}(B)$ to conclude $A \subset B$ or $B \subset A$. If we look back two pages, ← top of page! we can use the implications to realize that in order for $S \in \mathcal{S}(A \cup B) \Rightarrow \mathcal{S}(A) \cup \mathcal{S}(B)$, we need $S \subset A \cup B$ to imply $S \subset A$ or $S \subset B$. From there, we may be able to get to our conclusion. (It is delicate, however!)

Instead of taking that path, we might hope to slightly simplify our proof by turning the "or" statement in the conclusion ("$A \subset B$ or $B \subset A$") into an and statement, which may reduce the number of cases to consider. Instinctively, we may want to prove the contrapositive!

If $A \not\subset B$ and $B \not\subset A$, then $\mathcal{S}(A \cup B) \not\subset \mathcal{S}(A) \cup \mathcal{S}(B)$.

(*) Also means that $\mathcal{S}(A \cup B) \neq \mathcal{S}(A) \cup \mathcal{S}(B)$, the contrapositive of our original statement. The difference is now, we know exactly how to break equality!

Now, let's try. What does A not being a subset of B mean? We need to negate the definition of subset: there exists an $a \in A$ such that $a \notin B$. Similarly, B not being a subset of A means there exists a $b \in B$ such that $b \notin B$. If we start a proof by assuming $A \not\subseteq B$ and $B \not\subseteq A$, then we get the existence of the elements a and b above.



On the flip side, we want to show that $\mathcal{S}(A \cup B)$ is not a subset of $\mathcal{S}(A) \cup \mathcal{S}(B)$. Using the definition of subset (negated), we need to show that there exists an element $S \in \mathcal{S}(A \cup B)$ such that $S \notin \mathcal{S}(A) \cup \mathcal{S}(B)$.

In other words, $S \subseteq A \cup B$ but $S \not\subseteq A$ or $S \not\subseteq B$. So S needs to be in both A and B ... not just one or the other...

Our goal is now: construct S using the elements a and b above. If we can, we will have proven our claim.

Notice $\{a\} \subset A$, so picking $S = \{a\}$ won't work. ($S \subset A \cup B$ and $S \subset A$). However, if we add b...

$$S = \{a, b\}$$

we are in luck:

$$S = \{a, b\} \subset A \cup B$$

but

$$S = \{a, b\} \not\subset A \quad \text{since } b \notin A$$

and

$$S = \{a, b\} \not\subset B \quad \text{since } a \notin B.$$

Since we assumed $A \not\subset B$ and $B \not\subset A$, we were guaranteed the existence of a and b. That means we are guaranteed the existence of a set $S = \{a, b\}$, so we found a set $S \in \mathcal{P}(A \cup B)$ such that $S \notin \mathcal{P}(A) \cup \mathcal{P}(B)$.

Now we have all of the necessary elements to write a proof.

Problem 3

* we call such a g a right-inverse

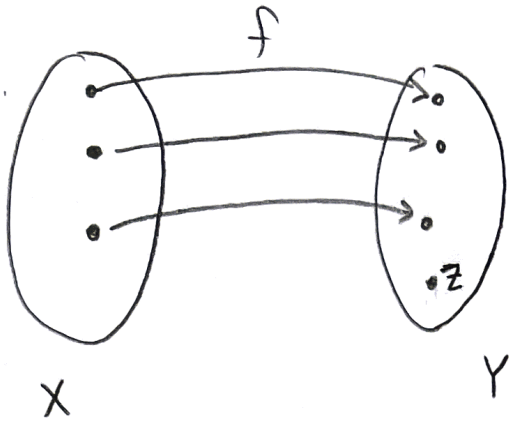
Let $f: X \to Y$ be a function. Prove that there exists a function $g: X \to Y$ such that $f \circ g = i_Y$ if and only if f is surjective.

(i_Y: identity map on Y (see Week #5))

Sketch / finding a proof:

First, does the statement seem true? If you have heard of a right-inverse, it may seem correct to you... but if you haven't, it may just look like a bunch of symbols and words we've seen recently...

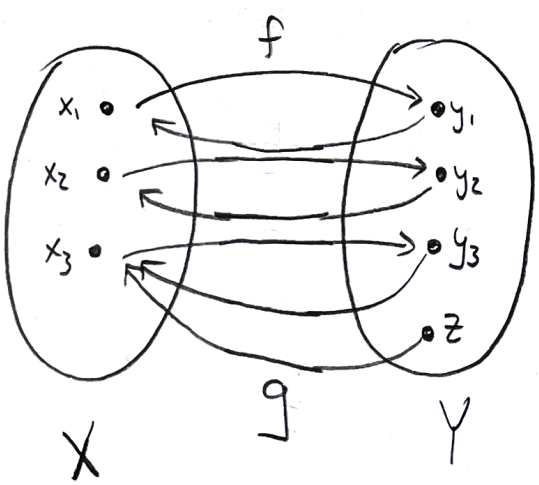
Let's try to pull the statement apart. Notice that the if and only if statement requires, in some sense, that f be surjective. This seems curious, and perhaps a good place to start pulling: why does f need to be surjective? Let's try to construct a function g such that $f \circ g = i_Y$, but that f is not surjective. (We should fail! But in our failure, hopefully we will see why we need such a condition on f.)



← $z \in Y$ is not in the image of X, in other words, z is not in $f(X)$, so f is not surjective.

Now, lets try to construct g so that $f \circ g = I_Y$.

There is a somewhat obvious candidate ... if we want $f \circ g(y) = i_Y(y) = y$ for all $y \in Y$, then g needs to send each $y \in Y$ to the element in X that f will map back to y. Since f doesn't map to z, we will need to pick something random for z ... (in fact... z may end up being the problem). Let's try:



Now, notice $f \circ g(y_1) = f(g(y_1)) = f(x_1) = y_1$, so that looks like

i_Y. The same works for y_2 and y_3. But ... something is funky with z:

$$f \circ g(z) = f(g(z)) = f(x_3) = y_3 \ \ldots$$

but $i_Y(z) = z \neq y_3$.

This is the problem: the element z that is not in $f(X)$ will never map back to itself! No matter what g we pick. This is our first observation: to construct g, we will need f to be surjective. This is why "f surjective $\Rightarrow$ there exists a $g: Y \to X$ such that $f \circ g = i_Y$" might be a reasonable thing to prove, and our observation says we will need to use surjectivity to construct g.

Now, before we try to prove this implication, let's think a bit about the other implication. How might we prove, "If there exists a function $g: Y \to X$ such that $f \circ g = i_Y$, then f is surjective"?

Does this implication look like anything we have seen before? If we look back in the Week #5 lecture notes, we will see that we did have a theorem in which we concluded a map is surjective:

THM 4 (2), second implication

$$f \circ g \text{ surjective} \Rightarrow f \text{ is surjective.}$$

In other words, if we can show $f \circ g$ is surjective, then by THM 4 (2) above, we can immediately conclude f is surjective! This is our second observation.

Now, with these two observations, we have strategies for approaching the proof of each implication.

Let's try showing if there exists $g: Y \to X$ such that $f \circ g = i_Y$, then f is surjective. We get to assume such a g exists ... and we really only need to show that $f \circ g = i_Y$ is surjective. By THM 4 (2), we can conclude f is surjective, assuming this is true. Let's prove $f \circ g$ is surjective.

Since $f \circ g = i_Y$, this will amount to showing that the identity map is surjective. In fact, you might suspect that the identity map is bijective, and you would be right. We haven't proven this though, so if we need it, we should show it:

- i_Y injective: Assume $y_1 \neq y_2$ and $y_1, y_2 \in Y$. Then $i_Y(y_1) = y_1 \neq y_2 = i_Y(y_2)$, by definition of the identity map. Thus, by definition of injective, we see that i_Y is injective.

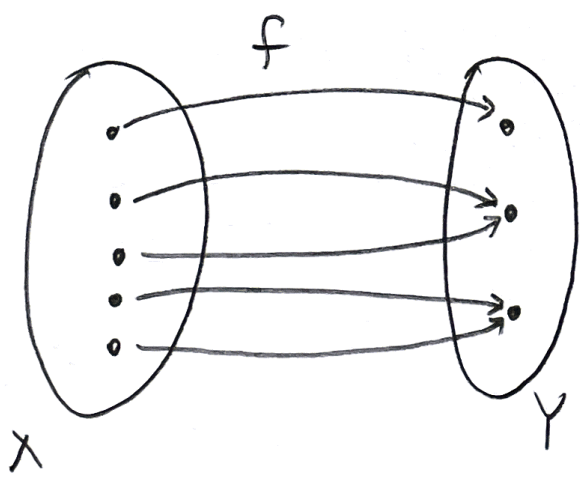
- i_Y surjective: Let $y \in Y$, the codomain of $i_Y : Y \to Y$. Is there an element in the domain Y such that i_Y sends it to y? Sure, $y \in Y$ will work: $i_Y(y) = y$. Since we picked $y \in Y$ arbitrarily in the beginning, this shows surjectivity by definition.

Hence, i_Y is bijective, being both injective and surjective. But this is not quite what we want to show. We want to show that $f \circ g$ is surjective. This won't be too hard since $f \circ g = i_Y$! We can pretty much recycle the surjectivity proof of i_Y. That gives us $f \circ g$ surjective, from which THM 4 (2)

allows us to conclude f is surjective.

Great. So that gives us everything we need to write a proof of that implication. Let's try to find all of the necessary pieces to get the other proof working. We need to show if f is surjective, then there exists a function $g: Y \longrightarrow X$ such that $f \circ g = i_Y$.

If we assume f is surjective, that means every $y \in Y$ has the property that $f(x)=y$ for some $x \in X$. We won't have the "z" problem from earlier. Let's draw this ... or at least something representative. We are not assuming f is injective, so let's not make it injective:



Our goal is to construct a g so that $f \circ g$ is i_Y.

To do this, we need, for all $y \in Y$,

$$f \circ g(y) = i_Y(y) = y.$$

Since $f \circ g(y) = f(g(y))$, and we want $f \circ g(y) = y$, then we need g to send $y \in Y$ to some $x \in X$ such that $f(x) = y$. (Does this look similar to a definition you know?) Basically, we have a choice. f might send more than one element to the same element in Y (see picture on previous page). We need to use one of these elements.

So, for each $y \in Y$, pick $x \in f^{-1}(\{y\})$, i.e. pick an x in the preimage of $\{y\}$. We know such an element exists since for every y, there is atleast one x such that $f(x) = y$ (definition of surjective!). This is how we are (crucially) using surjectivity of f.

Then, we get a function $g : Y \to X$ such that $g(y) = x$ implies $f(x) = y$. This gives us

all of the pieces we need! We can show that such a function $g: Y \to X$ has the property that $f \circ g = i_Y$, we just need to write this carefully.

Great. So hopefully these longer examples help you see how to prod and pull apart statements with the intent of proving them. It's not always easy, but hopefully you find it a satisfying process once you have a working proof. But as I mentioned on the first page of these notes, finding the proof is only step one. Carefully writing it is step two.

HOMEWORK

For your homework, provide cleaned-up, well-written proofs of Problems 2 and 3. Follow the writing guidelines in the email I sent to the class on Tuesday, July 27th, and the lecture notes from week #2. Make sure the first sentence of each proof does what it needs to do, make sure each line in your proof is justified or backed up by a definition and/or theorem, and do not add any extraneous or tangential lines in your proofs.

— More on next page! —

Homework Question 1

Complete Problem 2 using the contrapositive method.

Homework Question 2

Complete Problem 3.

- - - - - - - -

Problem 4

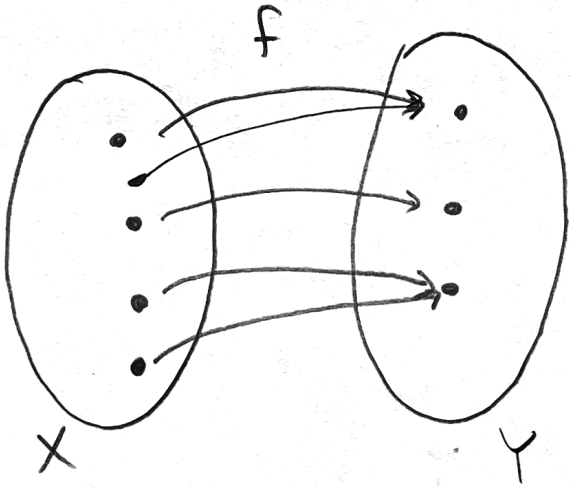
Construct a counterexample to the following statement.
If $f: X \longrightarrow Y$ and $g: Y \longrightarrow X$ are maps, and $f \circ g$ is surjective, the g is surjective.

Sketch / Finding a counterexample:

In problem 3, we recalled that $f \circ g$ surjective implies f is surjective. This is the content of theorem 4 from Week #5. In fact, in that same theorem, we show that if f and g are surjective, then so is the composition. The natural question here is whether the converse is true, i.e. does

$f \circ g$ being surjective imply f and g is surjective? The importance of the counterexample we are being asked to construct is that if it exists, it means this converse cannot be true! The process of constructing the counterexample gives us intuition as to why it cannot be true.

Our goal is this: construct an f and g so that $f \circ g$ is surjective, but g is not. Now, we know that if $f \circ g$ is surjective, it is necessary for f to be surjective. Let's start there.

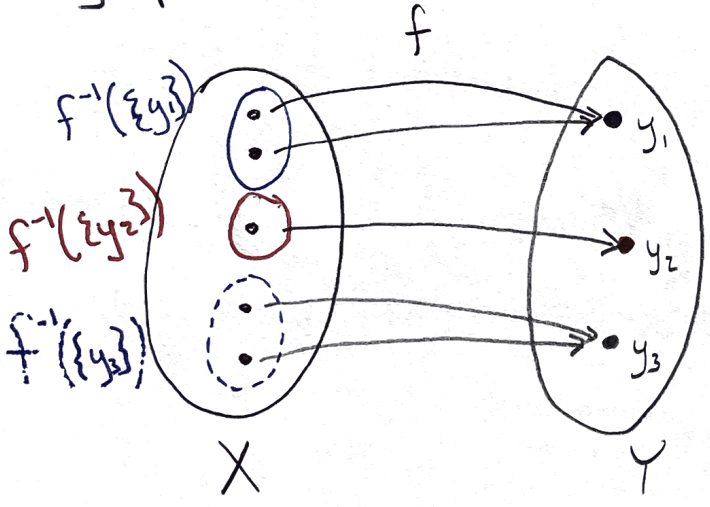


Notice, we do not need for f to be injective, so we intentionally make f not injective.

Now, we need a g that is not surjective. But we also want to pick g so that $f \circ g$ is surjective.

That means, for every $y \in Y$, we need to make sure there is an x in the preimage of $\{y\}$ that g maps to.

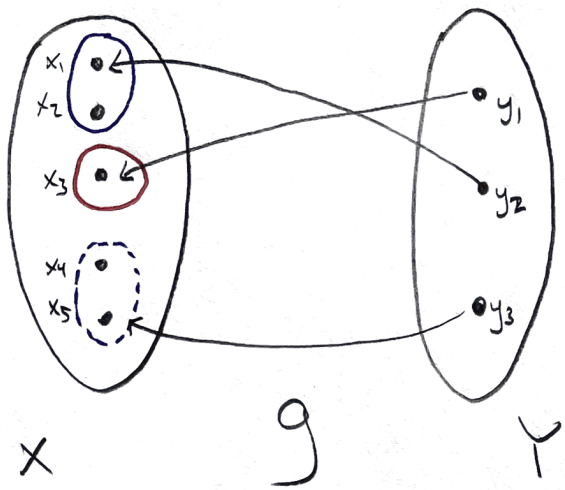
In other words, g must map into the preimage of every point in Y!



(Notice, this requirement is less restrictive than the requirement for g in Problem 3! We don't really need the composition to be the identity map on Y.)

For example, we could send y_2 to an element in $f^{-1}(\{y_1\})$, y_1 to an element in $f^{-1}(\{y_2\})$,

and y_3 to an element in $f^{-1}(\{y_3\})$.



Then, the composition will be surjective, but notice, g need not be surjective! If, however, there were exactly one element in the preimage of each $y \in Y$ under f (i.e. if f were injective!), we might construct a map g that is surjective ... which is precisely what we were trying to avoid. (In other words, it was important that we created an f that was not injective!)

So this tells us precisely how to construct g. Let's make all of this very concrete by giving the counterexample based on our construction above.

Counter example:

* went ahead and made them numbers instead of variables (not really necessary, but nice to have a concrete set).

Let $X = \{1, 2, 3, 4, 5\}$, and let $Y = \{1, 2, 3\}$. Define a map $f: X \longrightarrow Y$ as follows:

$$\begin{cases} f(1) = 1 \\ f(2) = 1 \\ f(3) = 2 \\ f(4) = 3 \\ f(5) = 3. \end{cases}$$

Now define a map $g: Y \longrightarrow X$ as follows:

$$\begin{cases} g(1) = 3 \\ g(2) = 1 \\ g(3) = 5. \end{cases}$$

Consider $f \circ g: Y \longrightarrow Y$.

$$f \circ g(1) = f(g(1)) = f(3) = 2$$

$$f \circ g(2) = f(g(2)) = f(1) = 1$$

$$f \circ g(3) = f(g(3)) = f(5) = 3.$$

Notice, $f \circ g$ is surjective: for any $y \in Y$, there exists a $\tilde{y} \in Y$ such that $f \circ g(\tilde{y}) = y$. However, g is not. $2 \notin g(Y)$.