Discrete Mathematics lecture notes 5-2

October 3, 2013

17. Learning to count

Let's make precise an intuitive notion that we've already encountered: The size of a set.

Notation 1. For $n \in \mathbb{N}$, let $[n]$ be the set $\{1, 2, \ldots, n\}$. In case that $n = 0$, note that $[0] = \emptyset$.

Definition 2. Let S be a set. S is *finite* if there exists an $n \in \mathbb{N}$ and a bijection $\varphi : [n] \xrightarrow{\sim} S$.[1] In this situation we say the *cardinality* (or *size*, or *order*, or...) of S is n, and write $|S| = n$.

There are two reasons for defining cardinality in terms of bijections. The deeper reason we will talk about next week, and involves our desire to talk about the size of *infinite sets*, where our notion of counting elements won't really make sense.[2] The second reason is that, while this may seem more cumbersome than a naive definition one might give, in fact it is actually the same thing.

Consider: Suppose I wanted to define "S is a finite set" to mean that I can write $S = \{a_1, a_2, \ldots, a_n\}$ for some finite n.[3] Consider the map $\varphi : [n] \to S : i \mapsto a_i$.[4] Note that φ is defined by our choice of how to number the elements of S. Then φ is injective, because we implicitly have not repeated an element of S in our list, and surjective because we have not left any element off of our list. Thus φ is a bijection, so the (somewhat more cumbersome, though more generalizable) definition in terms of the existence of a bijection is equivalent to our naive intuition. In other words, we're not wandering too far off the reservation, yet.

Continuing the pedantry: It is clear that the definition of S's being finite is *well-defined*, i.e., a set S is either finite (there exists a bijection $[n] \xrightarrow{\sim} S$) or it is not (there is no such bijection). On the other hand, it is not immediately clear that the definition of cardinality of a finite set is well-defined: Maybe there is some set S with a bijection $\varphi : [234] \xrightarrow{\sim} S$ and some other bijection $\psi : [235] \xrightarrow{\sim} S$. Would we say $|S| = 234$ and $|S| = 235$? Does that mean $234 = 235$? Hopefully this situation can never arise, but it will take a little work to prove as much.

18. Composition of functions

Recall that in Lecture Notes 5-1 we defined the notion of the *composite* of two relations: If R is a relation from A to B ($R \subseteq A \times B$) and S is a relation from B to C ($S \subseteq B \times C$), then

$$S \circ R := \{(a, c) \in A \times C | \exists\ b \in B \text{ s.th. } (a, b) \in R \text{ and } (b, c) \in S\}.$$

In the special case that R and S are both functions, the composite relation is still defined.

Definition 3. If $f : A \to B$ and $g : B \to C$ are functions, then the *composite of f and g* is the function $g \circ f : A \to C : x \mapsto g(f(x))$.

Exercise. Show that the composition of relations defined before specializes to the composition of functions.

We'd like to think of $\circ$ as a kind of algebraic operation, akin to multiplication in $\mathbb{Z}$ or $\mathbb{Z}/n$. There are some major differences, however. First, if $f : A \to B$ and $g : B \to C$, we can talk about $f \circ g : A \to C$, but the symbol $f \circ g$ *has no meaning*.[5] In the situation that $f : A \to B$ and $g : B \to A$, both of the compositions $f \circ g$ and $g \circ f$ exist, but it makes no sense to say that $f \circ g = g \circ f$.[6] Finally, even in the case that $f : A \to A$ and $g : A \to A$, we can ask whether $f \circ g = g \circ f$, but the answer is just "sometimes."

[1] The decoration $\sim$ on the arrow will indicate that the map in question is a bijection. Other standard notation: $f : A \hookrightarrow B$ indicates that f is an injection, and $g : X \twoheadrightarrow Y$ that g is a surjection.

[2] Spoiler alert: There are different sizes of infinity.

[3] Note: This is the same thing as saying "I can count the elements of S.

[4] More standard notation: We use the symbol $\to$ to indicate a function is going from the source (of the arrow) to the target (of the arrow), and the symbol $\mapsto$ to indicate where an *element* of the source gets sent. Thus the expressions $\varphi : [n] \to S : i \mapsto a_i$ is shorthand for "φ is a function with source $[n]$ and target S, such that $f(i) = a_i$ for all $i \in [n]$."

[5] Unless $C = A$.

[6] Unless $A = B$.

Exercise. Find a set A and two functions $f, g : A \to A$ such that $f \circ g \neq g \circ f$. Find two different functions $f', g' : A \to A$ such that $f \circ g = g \circ f$.

So composition of functions isn't always defined, let alone commutative. It is, however, associative.

Proposition 4. *Let $f : A \to B$, $g : B \to C$, and $h : C \to D$ be functions. Then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Proof. Follows immediately from the definition of composition. □

For A and B arbitrary sets, there isn't much we can say about the set of functions from A to B. In order to actually name any function (that is, identify the assignment $a \mapsto f(a) \in B$), we generally need to know *something* about the sets in question. There is one very important exception:

Definition 5. The *identity function of A* is the function $\mathrm{id}_A : A \to A : a \mapsto a$.

These identity functions[7] may not seem like much, but they are by far the most important functions that exist. It may take a while to believe this claim, but for the moment, let's show some of the things we can do with them.

Proposition 6. *Let $f : A \to B$ be any function. Then $\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A$.*

Definition 7. Let $f : A \to B$ and $g : B \to A$ be functions.

- g is a *left inverse to f* if $g \circ f = \mathrm{id}_A$.
- g is a *right inverse to f if* $f \circ g = \mathrm{id}_B$.
- g is an *inverse to f* if g is a left inverse and right inverse to f.

Note that if g is a left inverse to f, then f is a right inverse to g, etc.

Exercise. Show that if g and h are both inverses to f, then $g = h$. This allows us to speak of *the* inverse of f, written f^{-1}.

Exercise. Find a function $f : A \to B$ and functions $g, h : B \to A$ such that g and h are both left inverses to f, but $g \neq h$. Find another pair of functions $g', h' : B \to A$ that are both right inverses to f but $g \neq h$.

There is a close connection between these sided inverses and the special types of functions introduced in the previous lecture:

Proposition 8. *Let $f : A \to B$ and $g : B \to A$ be functions.*

(a) If g is a left inverse to f, then g is surjective and f is injective.

(b) If g is a right inverse to f, then g is injective and f is surjective.

(c) If $g = f^{-1}$, then both f and g are bijective.

Proof. First note that proving (a) will imply both (b) and (c): (b) is just (a) with the roles of f and g reversed, and (c) follows directly from (a) and (b) and the definition of bijectivity.

So suppose that $g \circ f = \mathrm{id}_A$. To show that g is surjective, for every element $a \in A$ we need to find some $b \in B$ such that $g(b) = a$. Consider $b = f(a)$. Then $g(b) = g(f(a)) = g \circ f(a) = \mathrm{id}_A(a) = a$, so g is surjective.

To show that f is injective, let $a, a' \in A$ be elements such that $f(a) = f(a')$. Then $g(f(a)) = g(f(a'))$, or $\mathrm{id}_A(a) = \mathrm{id}_A(a')$, or $a = a'$, as desired. □

In fact, the converse to these points are (almost) true as well:

Proposition 9. *Let $f : A \to B$ be a function.*

[7] Note: Every set has a *different* identity function, which is why we decorate the letters id to distinguish which one we're talking about.

(a) If f is injective and $A \neq \emptyset$, then there exists a left inverse to f.

(b) If f is surjective and we assume the Axiom of Choice, then there exists a right inverse to f.

(c) If f is bijective and we assume the Axiom of Choice, then there is a unique inverse to f.

Exercise. Why do I need to assume that $A \neq \emptyset$ in order for there to be a left inverse?

The condition in parts (b) and (c) that we assume this mysterious Axiom of Choice will have to remain that for the time being: Mysterious. Try to sketch a proof of (b) and see if you can pick out where you're forced to make a (surprisingly deep and far-reaching) assumption about the axioms that govern set theory. For now we will just assume that surjective functions always have right inverses.

As it happens, these notions we've introduced play nicely with composition.

Proposition 10. *Let $f : A \to B$ and $g : B \to C$ be functions.*

(a) If f and g are injective, then $g \circ f$ is injective.

(b) If f and g are surjective, then $g \circ f$ is surjective.

(c) If f and g are bijective, then $g \circ f$ is bijective.

Proof. We'll prove (c); the other two follow by the same logic. If f is bijective then f has a (unique) inverse $f^{-1} : B \to A$; similarly g has an inverse $g^{-1} : C \to B$. I claim that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. To check this, it's enough to show that $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \mathrm{id}_C$ and $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathrm{id}_A$. But both of these follow from the definition of the inverse and the fact that composition is associative. □

18. Detour: Functions to and from $\emptyset$

I make the following claim:

Theorem 11. *If A is any set, then there exists a unique function $f : \emptyset \to A$. If $A \neq \emptyset$, there is no function $A \to \emptyset$.*

Proof. This is really just a matter of thinking carefully about the definition of functions.

First, for A any set, I claim that $\emptyset \times A = \emptyset$. Why? Every element of $\emptyset \times A$ is an ordered pair (x, a) such that $x \in \emptyset$ and $a \in A$. But there is no $x \in \emptyset$, so there can be no such ordered pair. Similarly, $\emptyset \times A = \emptyset$.

Next, we consider the set of all *relations* from $\emptyset$ to A. Recall that this just means a subset $R \subseteq \emptyset \times A = \emptyset$, so there is precisely one, namely $R = \emptyset$. I claim that R is in fact a function. In order for this to be true, for all $x \in \emptyset$, there should exist a unique element $r \in R$ such that $(x, r) \in R = \emptyset$. This might seem like a contradiction, but remember that we're starting out assuming that $x \in \emptyset$, i.e., there is no such x. So we're testing the void hypothesis, and we see that the empty relation from $\emptyset$ to R is indeed a function.

Why doesn't this work the other way? Well, a function from A to $\emptyset$ is, first, a relation, and we've already seen that there is precisely one relation from A to $\emptyset$, namely $R = \emptyset$. Here's where the essential asymmetry of the definition of function comes into play: If $A \neq \emptyset$, there is some $a \in A$. Then there is no element $x \in \emptyset$ such that $(a, x) \in R$, because there is no $x \in \emptyset$. But this violates our definition that to *every* element of the domain, a function should assign a (unique) element of the range.[8] □

18. Finishing our counting education

Ok, let's return to the motivating question: Suppose that I have a particular set S and bijections $\varphi : [n] \xrightarrow{\sim} S$ and $\psi : [m] \xrightarrow{\sim} S$ for some $n, m \in \mathbb{N}$. Since bijections always have inverses,[9], I can look at $\psi^{-1} : S \xrightarrow{\sim} [m]$, another bijection. But then the composite $\psi^{-1} \circ \varphi : [n] \xrightarrow{\sim} [m]$ Will be a bijection. I claim this implies that $n = m$.

[8] In the language of category theory, $\emptyset$ is an *initial* object (i.e., there is a unique map *from* it *to* any other thing) in the category of sets. $\emptyset$ is not, however, *terminal* (there exists a unique map *to* it *from* any other set).

[9] Again, assuming the Axiom of Choice... though what about the case where you consider the empty set?

Theorem 12 (Pigeonhole Principle)**.** *Let n and m be natural numbers.*

(a) If $n < m$, there are no surjective maps $\varphi : [n] \hookrightarrow [m]$.

(b) If $n > m$. there are no injective maps $\psi : [n] \twoheadrightarrow [m]$.

(c) If $n \neq m$, there are no bijective maps $\eta : [n] \xrightarrow{\sim} [m]$.

Proof. We prove (c), because it's the one we need to show that set cardinality is well-defined. You should do the other two as exercises.

First, I claim that there is a *lexicographic ordering* on $\mathbb{N} \times \mathbb{N}$: Write $(a, b) \leq (a', b)$ if $a < a'$ or $a = a'$ and $b \leq b'$. I claim this is a total ordering.[10]

Moreover, this makes $\mathbb{N} \times \mathbb{N}$ into a well-ordered set: Let $S \subseteq \mathbb{N} \times \mathbb{N}$ be nonempty. For any pair $x = (a, b) \in S$, write $\pi_1(x) := a$ and $\pi_2(x) := b$. Then $\{\pi_1(x) | x \in S\}$ is a nonempty subset of $\mathbb{N}$, so it has a least element n. For that fixed n, let $T \subseteq S$ be $T := \{(n, b) \in S\}$. Then $\{\pi_2(y) | y \in T\}$ is a nonempty[11] subset of $\mathbb{N}$, so it too has a least element, say n. Then $(m, n) \in S$ is less than all other elements of S.

Ok, so $\mathbb{N} \times \mathbb{N}$ is well-ordered in this manner. Let us suppose for contradiction that there exist natural number n and m with $n \neq m$, together with a bijection $\varphi : [n] \xrightarrow{\sim} [m]$. Pick the smallest such pair (n, m). Note that $n \neq 0$, since there are no maps from a nonempty set to $[0] = \emptyset$, and similarly $m \neq 0$.

So let's focus on our bijection in our smallest counterexample: $\varphi : [n] \xrightarrow{\sim} [m]$. Define a new bijection $\alpha : [m] \to [m]$ by the rule:

$$\alpha(j) = \begin{cases} j & j < \alpha(n) \\ m & j = \alpha(n) \\ j - 1 & j > \alpha(n) \end{cases} .$$

Check that this is actually a bijection. This means $\alpha \circ \varphi : [n] \xrightarrow{\sim} [m]$ is also a bijection. By construction, $\alpha \circ \varphi(n) = m$, so if we *restrict* $\alpha \circ \varphi$ to $[n - 1]$ (i.e., throw away the point $n \in [n]$ and forget where it's sent in $[m]$), we get a bijection $[n - 1] \xrightarrow{\sim} [m - 1]$. But since $n, m > 0$, this is a different, smaller pair of unequal natural numbers, contrary with our assumption that (n, m) was the smallest pair where this was possible. Contradiction. □

[10] Check this!
[11] Why?