Functions and Cardinality Comparisons

Professor Cian Dorr

9th September 2022

New York University

How to write informal proofs: an example

Theorem. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. By the Axiom of Extensionality, it suffices to show that (a) $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$ and (b) $(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$.

(a) Suppose $x \in A \cap (B \cup C)$. Then $x \in A$, and either $x \in B$ or $x \in C$. In the first case, $x \in A \cap B$; in the second case, $x \in A \cap C$, so either way, $x \in (A \cap B) \cup (A \cap C)$.

(b) Suppose $x \in (A \cap B) \cup (A \cap C)$. Then there are two cases: either $x \in A \cap B$, or $x \in A \cap C$. In the first case, $x \in A$, and also $x \in B \cup C$ since $x \in B$, hence $x \in A \cap (B \cup C)$. In the second case, $x \in A$, and also $x \in B \cup C$ since $x \in C$, hence $x \in A \cap (B \cup C)$. So either way, we have $x \in A \cap (B \cup C)$.

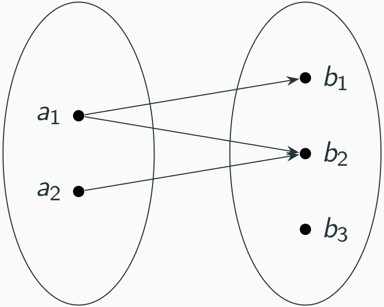
How to write informal proofs: another example

Theorem. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. By the Axiom of Extensionality, it suffices to show that for any x, $x \in A \cap (B \cup C)$ iff $x \in (A \cap B) \cup (A \cap C)$. So, consider an arbitrary x. $x \in A \cap (B \cup C)$ iff x belongs to both A and one of B or C. But this is the case iff either x is both in A and in B, or else both in A and in C. And that is the case iff $x \in (A \cap B) \cup (A \cap C)$.

Picturing relations

Let A be some two-membered set $\{a_1, a_2\}$ and B be a three-membered set $\{b_1, b_2, b_3\}$. Let R be the relation $\{\langle a_1, b_1\rangle, \langle a_2, b_2\rangle, \langle a_1, b_2\rangle\}$. Here are two ways of picturing this R:



	b_1	b_2	b_3
a_1	✓	✓	×
a_2	×	✓	×

Four properties of relations (again)

When R is a relation from A to B:

Definition

R is *serial* iff for every $x \in A$, there is some $y \in B$ such that Rxy.

Definition

R is *surjective* iff for every $y \in B$, there is some $x \in A$ such that Rxy.

Definition

R is *functional* [aka: a partial function] iff whenever Rxy and Rxy', $y = y'$.

Definition

R is *injective* iff whenever Rxy and $Rx'y$, $x = x'$.

Four properties of relations

These definitions immediately imply that for any relation R from A to B:

- R is serial iff R^{-1} is surjective.
- R is functional iff R^{-1} is injective.

It is also straightforward to prove (try it!) that when R is a relation from A to B and S is a relation from B to C:

- $S \circ R$ is [serial/surjective/functional/injective] if both R and S are.

Domain, codomain, range, corange

When R is a relation from A to B, we call A is *domain* and B its *codomain*.

The *range* of R is the set $\{y \in B \mid Rxy \text{ for some } x \in A\}$. This is not the same as the codomain (B) unless R is surjective.

The *corange* (not a standard term!) or "domain of definition" of R is the set $\{x \in A \mid Rxy \text{ for some } y \in B\}$. This is not the same as the domain (A) unless R is serial.

Functions

We have special names for some combinations of these properties.

Definition

A *function* from A to B is a relation that is both serial and functional.

Definition

An *injection* (one-to-one function) from A to B is an injective function from A to B.

Definition

A *surjection* (onto function) from A to B is a surjective function from A to B.

Definition

A *bijection* (one-to-one correspondence) from A to B is a function from A to B that is both injective and surjective.

Special notations for functions

Functions are so important that they deserve some special notation.

- $f : A \to B$ means 'f is a function from A to B'
- When $f : A \to B$, and $x \in A$, we write $f(x)$, or just fx, as short for 'the unique element y of B such that $\langle x, y \rangle \in f$'.
- Analogous to the set-builder notation for sets, we have "function builder" notation for functions. Examples:

$$[n \mapsto n^3 + n^2] : \mathbb{N} \to \mathbb{N}$$

$$[x \mapsto x\text{'s neighbor to the left}] : P \to P$$

$$\left[x \mapsto \begin{cases} x\text{'s dog} & \text{if } x \text{ has a dog} \\ \text{Lassie} & \text{otherwise} \end{cases}\right] : P \to D$$

Philosophical aside

Sameness of size

Algebra and arithmetic [are] the only sciences in which we can carry on a chain of reasoning to any degree of intricacy, and yet preserve a perfect exactness and certainty. We are possessed of a precise standard, by which we can judge of the equality and proportion of numbers; and according as they correspond or not to that standard, we determine their relations, without any possibility of error. When two numbers are so combined, as that the one has always a unit answering to every unit of the other, we pronounce them equal; and it is for want of such a standard of equality in extension, that geometry can scarce be esteemed a perfect and infallible science. (Hume, Treatise, 1.3.1)

Sameness of size

Inspired by Hume, we adopt the following definition of “equal in size”:

Definition

Sets A and B are *equinumerous* (notation: $A \sim B$) iff there is a bijection from A to B.

This has the following expected properties:

- $A \sim A$ for every set A (since id_A is a bijection from A to A.
- If $A \sim B$ then $B \sim A$ (since the converse of any bijection from A to B is a bijection from B to A).
- If $A \sim B$ and $B \sim C$ then $A \sim C$ (since the composition of any bijections from A to B and from B to C is a bijection from A to C).

At least as big as

If ‘A and B are equally big’ means that there’s a bijection from A to B, what should ‘B is at least as big as A’ mean? Well, here are two properties we’d expect this notion to have:

- If A and B are the same size, then B is at least as big as A.
- If $A \subseteq B$ then B is at least as big as A.

These suggest the following definition:

Definition

B is at least as big as A ($A \lesssim B$) iff A is equinumerous with some subset of B.

Or equivalently, bearing in mind that any bijection from A to a subset of B can also be considered as an injective function from A to B:

Alternative form

B is at least as big as A ($A \lesssim B$) iff there is an injective function from A to B.

Properties of ‘at least as big as’

Theorem

If $A \precsim B$ and $B \precsim C$, then $A \precsim C$.

Proof: suppose there's an injection f from A to B and an injection g from B to C. Then $g \circ f$ is an injection from A to C.

The following are much harder to show, and we won't prove them yet.

Schröder-Bernstein Theorem

If $A \precsim B$ and $B \precsim A$, then $A \sim B$.

Cardinal Comparability Theorem

For any sets A and B, either $A \precsim B$ or $B \precsim A$.

The proof of the Cardinal Comparability Theorem requires the *Axiom of Choice*, an additional axiom we haven't encountered yet.

Infinity

Definition

A is a *proper subset* of B ($A \subsetneq B$) $:=$ $A \subseteq B$ and not $A = B$.

Definition

A is *strictly smaller than* B ($A \lnsim B$) $:=$ $A \lesssim B$ and not $A \sim B$.

Given that we have $A \lesssim B$ whenever $A \subseteq B$, it is really tempting to assume that $A \lnsim B$ whenever $A \subsetneq B$, i.e. that any proper subset of a set is strictly smaller than it.

This is true for the empty set, for one-membered sets, for two-membered sets, for three-membered sets, ...

But it is false in general! Consider $\mathbb{N}$, the set of all natural numbers (including 0), and $\mathbb{N}^+$, the set of all *positive* natural numbers. We have $\mathbb{N}^+ \subsetneq \mathbb{N}$. Nevertheless, $\mathbb{N} \sim \mathbb{N}^+$, since the *successor* function which maps each number n to $n + 1$ is a bijection from $\mathbb{N}$ to $\mathbb{N}^+$.

Dedekind-infinity

Mathematician Richard Dedekind (1831–1916) proposed that this should be considered the defining feature of infinite sets.

Definition

A is *Dekind-infinite* $:=$ there is a bijection from A to a proper subset of A.

Equivalently: A is Dedekind-infinite iff there is a function from A to A that is injective but not surjective.

(We'll follow standard practice by reserving 'infinite' for a different definition we'll give later, which turns out to be equivalent to Dedekind-infinity given the Axiom of Choice.)

Facts about Dedekind-infinity

Fact 1

If A is Dedekind-infinite and $A \sim B$, then B is Dedekind-infinite.

Proof: suppose that A^- is a proper subset of A and $f : B \to A$ and $g : A \to A^-$ are bijections. Let $B^- = \{x \in B : fx \in A^-\}$, and let f^- be the restriction of f to B^-. Then f^- is a bijection from B^- to A^-, so $f^{-1} \circ g \circ f^-$ is a bijection from B^- to B.

Facts about Dedekind-infinity

Fact 2

If A is Dedekind-infinite and $A \subseteq B$, then B is Dedekind-infinite.

Proof: suppose that $f : A \to A$ is injective and a is not in its range. Define $g : B \to B$ as follows:

$$gx = \begin{cases} fx & \text{if } x \in A \\ x & \text{if } x \in B \backslash A \end{cases}$$

g is also injective. For suppose $gx = gy$. Then either $gx \in A$, in which case $gx = fx$ and $gy = fy$ so $x = y$ by the injectivity of f, or else $gx \in B \backslash A$, in which case $gx = x$ and $gy = y$ so $x = y$. But it is not surjective, since a is not in its range.

Fact 3

If A is Dedekind-infinite and $A \lesssim B$, then B is Dedekind-infinite.

Proof: immediate from facts 1 and 2.

Fact 4

If A is Dedekind-infinite and B isn't, then $B \lesssim A$.

Proof: from Fact 3 and the Cardinal Comparability Theorem.

Power Sets

The Power Set Axiom

For any set A, there is a set $\mathcal{P}A$ (or $\{B \mid B \subseteq A\}$) containing all and only the subsets of A.

(This is an obvious consequence of Naïve Comprehension, but not of Separation.)

$$\begin{aligned}
\mathcal{P}\varnothing &= \{\varnothing\} \\
\mathcal{P}\mathcal{P}\varnothing &= \{\varnothing, \{\varnothing\}\} \\
\mathcal{P}\mathcal{P}\mathcal{P}\varnothing &= \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\} \\
&\vdots
\end{aligned}$$

Fact: If A has n members, $\mathcal{P}A$ has 2^n members.

(We'll be able to prove this rigorously later when we introduce the natural numbers.)

Cantor's Theorem

Cantor's Theorem

For any set A, $A \lnsim \mathcal{P}A$: the power set of any set is strictly bigger than it.

This fits our observations about finite sets (since $2^n > n$ for all n).

But for infinite sets it's surprising: you might have thought that if a set is infinite, it's is *as big as can be*! But no: given our definition of 'bigger than', the powerset of an infinite set is a bigger infinite set.

Proving Cantor's Theorem

Cantor's Theorem

For any set A, $A \lnsim \mathcal{P}A$.

Proof: Clearly, $A \lesssim \mathcal{P}A$, since the function $[x \mapsto \{x\}] : A \to \mathcal{P}A$ is obviously injective. So it suffices to show that $A \nsim \mathcal{P}A$, i.e. that *no function from A to $\mathcal{P}A$ is a bijection*. We'll actually show that no function from A to $\mathcal{P}A$ is even a *surjection*.

Suppose that f is a function from A to $\mathcal{P}A$. Define f's "diagonal set" as:

$$D_f := \{x \in A \mid x \notin fx\}$$

Suppose for contradiction that there's an element $y \in A$ such that $fy = D_f$.

Then $y \in fy$ iff $y \in D_f$ (since $fy = D_f$)

Also, $y \in D_f$ iff $y \notin fy$ (since for any $x \in A$, $x \in D_f$ iff $x \notin fx$).

So, $y \in fy$ iff $y \notin fy$: Contradiction!

A note on the terminology of ‘diagonal set’

When we have a relation R from A to A, $\operatorname{diag} R$ is the set $\{x \in A \mid Rxx\}$. It’s so-called because we can determine what’s in it by representing R as a grid of ticks and crosses, and looking down the diagonal of the grid.

$$\begin{array}{c|cccc} & a & b & c & d \\ \hline a & \checkmark & \checkmark & \times & \checkmark \\ b & \times & \times & \times & \checkmark \\ c & \checkmark & \times & \checkmark & \checkmark \\ d & \checkmark & \times & \checkmark & \times \end{array} \quad \Rightarrow \quad \begin{array}{c} \boxed{\checkmark} \\ \boxtimes \\ \boxed{\checkmark} \\ \boxtimes \end{array}$$

But what does that use of ‘diagonal’ have to do with functions from a set to its powerset?

Currying and uncurrying

Well, it turns out there's a very close connection between (a) *relations from A to B* and (b) *functions from A to $\mathcal{P}B$*.

- Any relation R from A to B determines a function $\operatorname{curry} R$ from A to $\mathcal{P}B$, namely $[x \mapsto \{y \in B \mid Rxy\}]$.
- Any function f from A to $\mathcal{P}B$ determines a relation $\operatorname{uncurry} f$ from A to B, namely $\{\langle x, y\rangle \mid y \in fx\}$.

These operations are inverses:

$$\begin{aligned}
\operatorname{curry}(\operatorname{uncurry} f) &= [x \mapsto \{y \mid \langle x, y\rangle \in \operatorname{uncurry} f\}] \\
&= [x \mapsto \{y \mid y \in fx\}] = [x \mapsto fx] = f \\
\operatorname{uncurry}(\operatorname{curry} R) &= \{\langle x, y\rangle \mid y \in (\operatorname{curry} R)x\} \\
&= \{\langle x, y\rangle \mid y \in \{z \mid Rxz\}\} = \{\langle x, y\rangle \mid Rxy\} = R
\end{aligned}$$

Why we called D_f the "diagonal set"

Given our $f : A \to \mathcal{P}A$, we have:

$$\begin{aligned} \operatorname{curry} f &:= \{\langle x, y\rangle \in A \times A \mid y \in fx\} \\ \operatorname{diag}(\operatorname{curry} f) &:= \{x \in A \mid \langle x, x\rangle \in \operatorname{curry} f\} = \{x \in A \mid x \in fx\} \end{aligned}$$

and so

$$A \setminus \operatorname{diag}(\operatorname{curry} f) := \{x \in A \mid x \notin fx\} = D_f$$