MATH 301

INTRODUCTION TO PROOFS

Sina Hazratpour
Johns Hopkins University

Fall 2021

- Relations
- Functions

Relevant sections of the textbook

- Chapter 3
- Chapter 5

Relations

Definition

A *(binary) relation* R on sets A and B is a two-variable predicate $R(x, y)$ where $x \in A$ and $y \in B$.

Remark

In mathematics, we often use infix notation, writing $a\ R\ b$ instead of $R(a, b)$, e.g. $a = b$, $a \leqslant b$, $f \pitchfork g$, etc.

Definition

An *extension* (aka graph) of a relation R on sets A and B is a subset $[R]$ of $A \times B$ consisting of the pairs (a, b) where aRb.

Exercise

Prove that any subset of $A \times B$ is obtained as an extension of some relation on A and B.

Examples of relations

1. The relation on days on the calendar, given by x and y fall on the same day of the week.
2. The relation on vegetable produce, given by price of x is less than price of y.
3. The relation on people currently alive on the planet, given by x and y have the same home address.
4. The relation on people in the world, given by x is a brother of y.
5. The relation on people in the world, given by person x is influenced by person y.
6. The relation on lines on a 2-dim plane, given by line l and line m are parallel to each other.
7. The relation on points and lines on a 2-dim plane, given by point p is on line l.

Equivalence relation

Definition

A binary relation R on a domain A is an equivalence relation if it has the following three properties:

(reflexivity) *aRa, for every a in A.*

(transitivity) *If aRb and bRc, then aRc, for every a, b, and c in A.*

(symmetry) *If aRb then bRa, for every a and b in A.*

Which of the relations of the previous slide are

- reflexive?
- transitive?
- symmetric?

Partial order

Definition

A binary relation R on a domain A is a partial order if it has the following three properties:

(reflexivity) *aRa, for every a in A.*

(transitivity) *If aRb and bRc, then aRc, for every a, b, and c in A.*

(antisymmetry) *If aRb and bRa then a = b, for every a and b in A.*

Which of the relations of the previous slide are

- anti-symmetric?
- a partial order?

Observe that the relation of strict inequality between integers is not a partial order since it is not reflexive.

The following are all examples of partially ordered sets (aka posets):

- $\leqslant$ on the natural numbers;
- $\leqslant$ on the integers;
- $\leqslant$ on the rational numbers;
- $\leqslant$ on the real numbers.

Because of the great many uses and ubiquity of the above examples of partial order in mathematics, we simply use the symbol $\leqslant$ for a general partial order R.

But keep in mind that $\leqslant$ is only a symbol and it is only meant to be suggestive; it can have unexpected interpretations as well. For example, the $\geqslant$ relation on any of these domains is also a partial order, and can interpret the $\leqslant$ symbol just as well.

Total order

Definition

A partial order R on a domain A is a total order (also called a linear order) if it also has the following property: for every a and b in A, either aRb or bRa.

Example

Show that the set

$$\overline{n} = \{\overline{0}, \overline{1}, \cdots, \overline{n-1}\},$$

which we studied before, has a partial order given by subset relation. Show that this order is total.

Associated graph of a relation

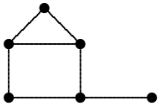
Suppose a set A comes equipped with a relation R. We can associate a directed graph (aka a digraph) with vertex set A and with an ordered pair $(a, b) \in A \times A$ being an edge precisely when aRb.

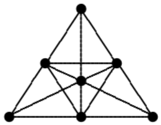
Exercise

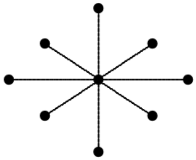
Express the conditions of reflexivity, transitivity, symmetry, antisymmetry, and totality in terms of familiar connectivity conditions on the associated graph.

Exercise

If the following graphs are the associated graphs of certain relations, what facts about those relations can we infer?







Exercise (Partial order on a power set)

There is a partial order on a power set $\mathcal{P}(X)$ of a set X given by the subset relation: Check that all the axioms of partial order are satisfied. Show that this partial order is not total.

$$\begin{tikzcd}
 & \{x,y,z\} & \\
\{x,y\} \arrow[ur] & \{x,z\} \arrow[u] & \{y,z\} \arrow[ul] \\
\{x\} \arrow[u] \arrow[ur, dotted] & \{y\} \arrow[ul] \arrow[ur] & \{z\} \arrow[u] \arrow[ul, dotted] \\
 & \emptyset \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

In fact we can recover the partial order of $\mathcal{P}(X)$ simply from the intersection (or equivalently the union) operation.

For subsets A, B of X, define

$$A \leqslant B \iff A \cap B = A$$

Exercise

Show that $\leqslant$ is a partial order relation, and it agrees with the subset relation.

Definition

A non-empty partially ordered set $(S, \leqslant)$ is filtered (or is said to be a filtered set) if for each $a, b \in S$, there is a element c such that $a \leqslant c$ and $b \leqslant c$.

Remark

Every total order is a filtered.

Example

The powerset $\mathcal{P}(X)$ with the subset relation is filtered.

Exercise

Show that for a poset P the set of filtered subsets of P is again filtered.

Minimum and maximum

Definition

We say an element a of a poset P is a minimum (aka a least element) for P if it is less than or equal to any other element, that is

$$\forall x \in P\ (a \leqslant x)$$

Dually, we say an element a of a poset P is a maximum (aka a greatest element) for P if it is greater than or equal to any other element, that is

$$\forall x \in P\ (x \leqslant a)$$

Example

- *In* $(\mathbb{N}, \leqslant)$, 0 *is a minimum; there is no maximum.*
- *Let* $n \in \mathbb{N}$ *with* $n > 0$. *Then* $\underline{0}$ *is a least element of* $(\underline{n}, \leqslant)$, *and* $\underline{n-1}$ *is a greatest element.*
- $(\mathbb{Z}, \leqslant)$ *has no maximum or minimum.*
- *The interval* $((0, 1], \leqslant)$ *has a maximum but not a minimum.*

Definition

We say that an element is ***minimal*** *for a partial order if no element is less than it. Dually, we say that an element y is* ***maximal*** *for a partial order if no element is greater than it.*

Our logical idea of function

A function f from a set X to a set Y is a specification of a unique element $f(x) \in Y$ for each $x \in X$. We write $f\colon X \to Y$ to denote the assertion that f is a function with domain X and codomain Y.

To describe a particular function, one must specify

- its domain,
- its codomain, and
- the effect of function upon a typical ("variable") element of its domain.

For instance the "squaring" function on the set of real numbers is specified in either of the following ways:

1. $f\colon \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$ for every real number x, or
2. $x \mapsto x^2 \colon \mathbb{R} \to \mathbb{R}$,
3. $\lambda x.x^2 \colon \mathbb{R} \to \mathbb{R}$.

How to define a function? (I)

The simplest way to define a function is to give its value at every x with an explicit well-defined expression.

Example

- *Let* $f : \mathbb{N} \to \mathbb{N}$ *be the function defined by* $f(n) = n + 1$*, that is* $f = \lambda n.n + 1$.
- *Let* $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be the function defined by* $g(x, y) = x^2 + y^2$.
- *Let* $p : \mathbb{N} \to \mathbb{N}$ *be the function defined by* $p(n)$ = *the largest prime number less than or equal to* n.
- *The assignment to each real number the greatest integer less than or equal to it. We call this function the floor function. We denote this function by* $\lfloor - \rfloor : \mathbb{R} \to \mathbb{Z}$.
- *The assignment to each real number the least integer greater than or equal to it. We call this function the ceiling function. We denote this function by* $\lceil - \rceil : \mathbb{R} \to \mathbb{Z}$.

Some functions on power sets

Example

- $\lambda x.\{x\} \colon X \to \mathcal{P}(X)$. *We sometimes denote this function by* $\{-\}$.
- $\lambda A. \bigcup_{a \in A} a \colon \mathcal{P}(\mathcal{P}(X)) \to \mathcal{P}(X)$.

How to define a function? (II)

It is sometimes convenient to define a function using different specifications for different elements of the domain.

Example

The absolute value function $|-| : \mathbb{R} \to \mathbb{R}$*, defined for* $x \in \mathbb{R}$

$$|x| = \begin{cases} x & \textit{if } x \geqslant 0 \\ -x & \textit{if } x \leqslant 0 \end{cases}$$

When specifying a function $f : X \to Y$ by cases, it is important that the conditions be:

- exhaustive: given $x \in X$, at least one of the conditions on X must hold; and
- compatible: if any $x \in X$ satisfies more than one condition, the specified value must be the same no matter which condition is picked.

Characteristic functions

Definition

Let X be a set and let $U \subseteq X$. The characteristic function of U in X is the function $\chi_U \colon X \to \{0, 1\}$ defined by

$$\chi_U(a) = \begin{cases} 1 & \text{if } a \in U \\ 0 & \text{if } a \notin U \end{cases}$$

Example

$\chi_E : \mathbb{N} \to \{0,1\}$ *is the function defined by*

$$\chi_E(n) = \begin{cases} 0 & \textit{if } n \textit{ is even} \\ 1 & \textit{if } n \textit{ is odd.} \end{cases}$$

$\chi_{\mathbb{Q}} : \mathbb{R} \to \{0,1\}$ *is the function defined by*

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 0 & \textit{if } x \textit{ is rational} \\ 1 & \textit{if } x \textit{ is irrational.} \end{cases}$$

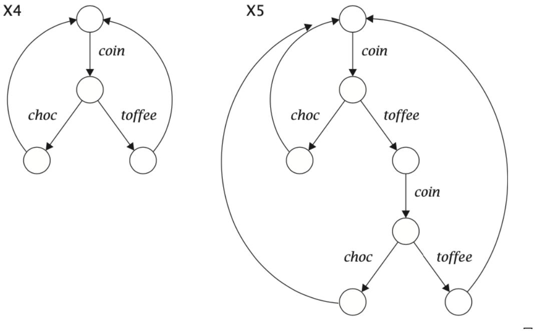
Try to draw the graph of the second function, or at least try to imagine it in your mind.

Exercise

Show that

1. $\chi_{U\cap V} = \chi_U \chi_V$
2. $\chi_{U\cap V} = \chi_U + \chi_V - \chi_U \chi_V$
3. $\chi_{U^c} = 1 - \chi_U$

Our mechanistic idea of function



Functions as machines

We might think of a function as a *machine* which, when given an *input*, produces an *output*. This “machine” is defined by saying what the possible inputs and outputs are, and then providing a list of instructions (an *algorithm*) for the machine to follow, which on any input produces an output—and, moreover, if fed the same input, the machine always produces the same

Warning

Our algorithmic idea of function implies that functions are computable in some sense. Note that this idea is at odds with a view of functions as well-formed logical expressions.
For example, concerning the characteristic function $\chi_{\mathbb{Q}}$, it is not at all clear what it means to be presented with a real number as input, let alone whether it is possible to determine, algorithmically, whether such a number is rational or not.

It is much harder to make formal what is meant by an "algorithm". This was first done by Alan Turing and Alonzo Church.

Equality of functions

Definition (function extensionality)

Functions $f\colon X \to Y$ and $g\colon X \to Y$ are equal if and only if the sentence

$$\forall x \in X\, f(x) = g(x)$$

is true.

Exercise

Show that for any set A there is a unique function $\emptyset \to A$.

Compositionality of functions

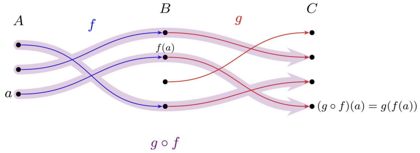
For any set X, we can define a function $\mathrm{id}\colon X \to X$ by letting $\mathrm{id}(x)$ to be the same as x. This function is called the identity function on X.

More interestingly, let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions. We can define a new function $k\colon X \to Z$ by letting

$$k(x) =_{\mathrm{def}} g(f(x))$$

The function k is called the composition of f and g which we also call "f composed with g" (or "g after f") and which we denote by $g \circ f$.

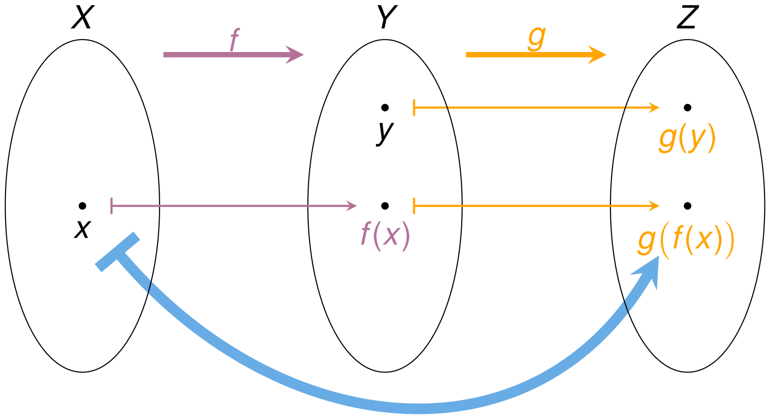
$$\begin{tikzcd} & Y \arrow[dr, "g"] & \\ X \arrow[ur, "f"] \arrow[rr, "g\circ f"'] & & Z \end{tikzcd}$$



The order of composition

The order of composition is somewhat confusing; the syntactic order does not match the diagrammatic order. In the diagram above, f appears to the left of g while in the syntactic expression of composition $g \circ f$, the function f appears appears on the right.

Nevertheless, they both mean the same thing: in order to evaluate the expression $g(f(x))$ you first evaluate f on input x, and then evaluate g. The function g waits for the the result $f(x)$ of application of f to the input x and once that is available, g applies to the value $f(x)$.



$$\lambda y.g(y) \circ \lambda x.f(x) = \lambda x.g[f(x)/y]$$

$$\lambda y.log_2 y \circ \lambda x.2^x = \lambda x.log_2 y\,[2^x/y] = log_2 2^x = x$$

The composition of function introduced above has two important properties:

unitality for any function $f\colon X \to Y$, we have $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_Y \circ f = f$.

associativity for any functions $f\colon W \to X$, $g\colon X \to Y$ and $h\colon Y \to Z$, we have

$$h \circ (g \circ f) = (h \circ g) \circ f\,.$$

Constant functions

Definition

We say a function $f\colon X \to Y$ is constant if for all $x, x' \in X$ we have $f(x) = f(x')$.

Exercise

Show that the identity function $\mathrm{id}\colon \emptyset \to \emptyset$ is constant.

Exercise

Suppose $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. Show that if either f or g is constant then the composition $g \circ f$ is constant.

Commuting diagrams of functions

We say a square

$$\begin{tikzcd}
A \arrow[r, "f"] & B \arrow[d, "g"] \\
C \arrow[u, "h"] \arrow[r, "k"'] & D
\end{tikzcd}$$

of sets and functions commutes if

$$g \circ f \circ h = k$$

Commuting diagrams of functions

We say a square

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "h"'] & B \arrow[d, "g"] \\
C \arrow[r, "k"'] & D
\end{tikzcd}$$

of sets and functions commutes if

$$g \circ f = k \circ h$$

Functions and relations

Functions can be seen as a special kind of relations.

Definition

A binary relation $R(x, y)$ on A and B is functional if for every x in A there exists a unique y in B such that $R(x, y)$. We can express this formally by the following sentence

$$\left(\forall x \exists y R(x, y)\right) \wedge \left(\forall x \forall y \forall z (R(x, y) \wedge R(x, z) \Rightarrow y = z)\right)$$

If R is a functional relation, we can define a function $f_R \colon X \to Y$ by setting $f_R(x)$ to be equal to the unique y in B such that $R(x, y)$. Conversely, it is not hard to see that if $f \colon X \to Y$ is any function, the relation $R_f(x, y)$ defined by $f(x) = y$ is a functional relation.

For any function $f\colon X \to Y$, we define as subset of $X \times Y$ known as the graph of f.

$$\mathbf{Gr}(f) = \{(x, y) \mid f(x) = y\}$$

Define functions h, i, and p as follows:

$$h = \lambda x.(x, f(x)) \tag{1}$$

$$i = \lambda (x, y).(x, y) \tag{2}$$

$$p = \lambda (x, y).y \tag{3}$$

Exercise

Show that the functions f, h, i, and p fit into the following square of sets and functions commutes:

$$\begin{tikzcd}
\mathbf{Gr}(f) \arrow[r, "i"] & X \times Y \arrow[d, "p"] \\
X \arrow[u, "h"] \arrow[r, "f"'] & Y
\end{tikzcd}$$

Composition of relations

Given a relation R on X and Y and a relation S on Y and Z we can compose them to get a relation $S \circ R$ on X and Z defined as follows:

$$x(S \circ R)z \iff \exists y \in Y\,(xRy \wedge yRz)$$

Exercise

Let B be the "brothership" relation (xBy means x is a brother of y) and S be the "sistership" relation. Show that the composite relation $S \circ B$ is not equivalent to B.

Exercise

- *Prove that if both R and S are partial orders then $S \circ R$ is a partial order.*
- *Prove that if both R and S are equivalence relations then $S \circ R$ is an equivalence relation.*

Composition of functions from compositions of relations

Theorem

Suppose $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. Consider the corresponding relations R_f and R_g. The relation corresponding to the composite function $g \circ f$ is equivalent to the composite relations $R_g \circ R_f$, that is,

$$\forall x \in X \forall z \in Z \left(x \, R_{g \circ f} \, z \iff x \, (R_g \circ R_f) \, z\right)$$

Isomorphisms of sets

Definition

An isomorphism between two sets X and Y is a pair of function

$$f\colon X \to Y \text{ and } g\colon Y \to X$$

such that $g \circ f = \mathrm{id}_X$, and $f \circ g = \mathrm{id}_Y$.

We can think of functions f and g above as no data-loss “processes”, e.g. conversion of files to different format without data being lost.

Definition

The sets X and Y are said to be isomorphic in case there exists an isomorphism between them. In this case, we use the notation $X \cong Y$.

Exercise

Show that for any set A, it is isomorphic to $\emptyset$ if and only if A does not have any elements. Can you prove this without the LEM?

Previously, we defined the cartesian product $A \times B$ of two sets A and B to consists of all the pairs (a, b) where $a \in A$ and $b \in B$. Now, we show that if we have more two sets the order of forming products does not matter.

Exercise

1. *For all sets A, B, C we have*

$$(A \times B) \times C \cong (A \times B) \times C$$

For this reason, we use $A \times B \times C$ to denote either sets.

Exercise

Show that two finite sets are isomorphic if and only if they have the same number of elements.

Exercise

Show that for any function $f\colon X \to Y$, we have

$$\mathbf{Gr}(f) \cong X\,.$$

A remark on disjoint unions

We introduced the operation of taking disjoint union of two sets as follows:

$$A \sqcup B = \{\mathrm{inl}(x) \mid x \in A\} \cup \{\mathrm{inr}(x) \mid x \in B\}$$

Exercise

Show that

$$A \sqcup B \cong (\{0\} \times A) \cup (\{1\} \times B)$$

Inspired by this fact we define the disjoint union of a family $\{A_i \mid i \in I\}$ of sets to be

$$\bigsqcup_{i \in I} A_i = \bigcup_{i \in I} \{i\} \times A_i \,.$$

An element of $\bigsqcup_{i \in I} A_i$ is a pair (i, a) where $i \in I$ and $a \in A_i$.

Arithmetic of sets

We define the operation of addition on sets as follows: For sets X and Y let the sum $X + Y$ be defined by their disjoint union $X \sqcup Y$.

Exercise

1. *Show that the addition operation on sets is both commutative and associative.*
2. *Show that the empty set is the unit (aka neutral element) of addition of sets.*

Exercise

Show that $\underline{m} + \underline{n} \cong \underline{m + n}$ *for all natural numbers* m *and* n.

Exercise

1. *Show that if S and S' are isomorphic, then for all sets X, we have $X + S \cong X + S'$.*
2. *Prove that for any singleton S, we have $\mathbb{N} + S \cong \mathbb{N}$.*

Sometimes, when the context precludes risk of confusion, we use the notation 1 for any singleton set. Therefore, we can simplify the last statement in above to

$$\mathbb{N} + 1 \cong \mathbb{N}.$$

Definition

- *A retract (aka left inverse) of function $f\colon A \to B$ is a morphism $r\colon B \to A$ such that $r \circ f = \mathrm{id}_A$. In this case we also say A is a retract of B.*
- *A section (aka right inverse) of function $f\colon A \to B$ is a morphism $s\colon B \to A$ such that $f \circ s = \mathrm{id}_B$.*

$$\begin{tikzcd}
 & A \arrow[rr, "\mathrm{id}_A"] \arrow[dr, "f"'] & & A \\
B \arrow[ur, "s"] \arrow[rr, "\mathrm{id}_B"'] & & B \arrow[ur, "r"'] &
\end{tikzcd}$$

Example

- *The circle is a retract of punctured disk.*
- *The maps from the infinite helix to the circle has a section, but no continuous section.*

Injections

Definition

A function $f\colon X \to Y$ is injective (or one-to-one) if

$$\forall a, b \in X,\ f(a) = f(b) \Rightarrow a = b$$

An injective function is said to be an injection.

Proposition

Let $f : X \to Y$ be a function. If f is injective and X is inhabited, then f has a retract.

Proof.

Suppose that f is injective and X is inhabited. Since X is inhabited, we get always fix an element of it, say $x_0 \in X$. Now, define $r \colon Y \to X$ as follows.

$$r(y) = \begin{cases} x & \text{if } y = f(x) \text{ for some } x \in X \\ x_0 & \text{otherwise} \end{cases}$$

Note that r is well-defined since if for some y, the there are elements x and x' such that $y = f(x) = f(x')$, then, by injectivity of f, we have $x = x'$, and therefore, the value of r is uniquely determined.

To see that r is a retract of f, let $x \in X$. Letting $y = f(x)$, we see that y falls into the first case in the specification of r, so that $r(f(x)) = g(y) = a$ for some $a \in X$ for which $y = f(a)$. But, $f(x) = y = f(a)$, and by injectivity of f we have $x = a$. Therefore, for every $x \in X$,

$$r(f(x)) = x = \mathrm{id}_X(x)\,.$$

Was this proof constructive?

Surjections

Definition

A function $f\colon X \to Y$ is surjective (aka onto) if

$$\forall y \in Y,\ \exists x \in X,\ f(x) = y$$

holds. A surjective function is said to be a surjection.

Proposition

Let $f : X \to Y$ be a function. If f is injective and X is inhabited, then f has a retract.

A function $f\colon X \to Y$ induces a function

$$\mathcal{P}(f)\colon \mathcal{P}(Y) \to \mathcal{P}(X)$$

defined by

$$\mathcal{P}(f)(S) = \{x \in X \mid f(x) \in S\}$$

for any subset S of Y. Note that

$$\mathcal{P}(\mathrm{id}_X) = \mathrm{id}_{\mathcal{P}(X)}$$

Suppose $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. We prove that

$$\mathcal{P}(f) \circ \mathcal{P}(g) = \mathcal{P}(g \circ f)\,.$$

Recall that in order to prove equality of functions we need to use function extensionality.

Suppose T is a subset of Z. Then

$$\begin{aligned}
\mathcal{P}(f) \circ \mathcal{P}(g) T &= \mathcal{P}(f)\,\{y \in Y \mid g(y) \in T\} \\
&= \big\{x \in X \mid f(x) \in \{y \in Y \mid g(y) \in T\}\big\} \\
&= \{x \in X \mid g(f(x)) \in T\} \\
&= \mathcal{P}(g \circ f) T
\end{aligned}$$

Fibres

Definition

For a function $f : X \to Y$, and an element $y \in Y$, the subset

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}$$

of X is called the fibre of f at y and also the pre-image of y under f.

Example

Consider the function $\lfloor - \rfloor : \mathbb{R} \to \mathbb{Z}$ which takes a real number to the greatest integer less than it. What are the fibres

- $\lfloor - \rfloor^{-1}(0)$?
- $\lfloor - \rfloor^{-1}(\lfloor \pi \rfloor)$?

The operation of taking fibres of a function is itself a function. More specifically, given a function f, taking fibres of f at different elements $y \in Y$ as a function is equal to the composite

$$Y \xrightarrow{\{-\}} \mathcal{P}(Y) \xrightarrow{\mathcal{P}(f)} \mathcal{P}(X)\,,$$

that is for all $y \in Y$,

$$f^{-1}(y) = \mathcal{P}(f)\{y\}$$

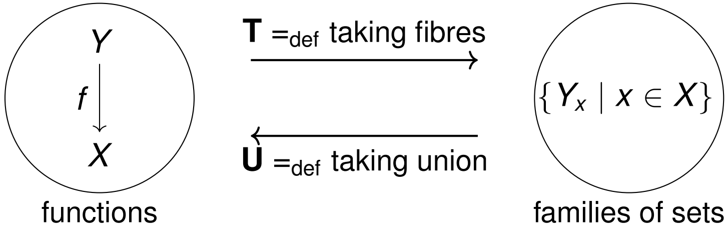
Exercise

Consider the family $\{f^{-1}(y) \mid y \in Y\}$. Show that all members of this family are mutually disjoint, and that their union is fact X.

$$\coprod_{y \in Y} f^{-1}(y) \cong \bigcup_{y \in Y} f^{-1}(y) = X$$

As the last exercise suggests, we can associate to every function a family of sets given by fibres of that function at different elements of the codomain.

Interestingly, we have the reverse association too: to a family $\{ Y_x \mid x \in X \}$ we associate a function as follows: let the domain to be the disjoint union $\bigsqcup\limits_{x \in X} Y_x$ and let the codomain be X. The associated function $p \colon \{ Y_x \mid x \in X \} \to X$ takes an element $\mathrm{in}(x) \in \bigsqcup\limits_{x \in X} Y_x$ to $x \in X$.



$Y \xrightarrow{f} X$

$\mathbf{T} =_{\mathrm{def}}$ taking fibres

$\mathbf{U} =_{\mathrm{def}}$ taking union

$\{ Y_x \mid x \in X \}$

functions

families of sets

The set of functions

Suppose X and Y are sets. We can define a new set consisting of all the functions from X to Y. We denote this set by Y^X. Explicitly,

$$Y^X = \{f\colon X \to Y\} \cong \{R \subset X \times Y \mid R \text{ is a functional relation}\}$$

Exercise

Suppose X is a finite set with m elements and Suppose Y is a finite set with n elements. Then the set Y^X has n^m elements.

The set of functions behaves like exponentials

Proposition

Suppose X, Y, Z are sets. We have

- $X^{\emptyset} \cong 1$
- $\emptyset^{X} \cong 1$ *if and only if* $X = \emptyset$. *In particular* $\emptyset^{\emptyset} \cong 1$.
- $(X^{Y})^{Z} \cong X^{Y \times Z}$.
- $X^{Y+Z} \cong X^{Y} \times X^{Z}$

Let Ω be a set with two elements, for instance $\{\top, \bot\}$. We show that

$$\Omega^X \cong \mathcal{P}(X)$$

that is the power set of X is isomorphic to the set of functions from X to Ω. To this end we construct two functions f and g and prove that they are inverse of each other. The function $f\colon \Omega^X \to \mathcal{P}(X)$ is defined as $\lambda(\varphi : \Omega^X).\{x \in X \mid \varphi(x) = \top\}$.

The function $g\colon \mathcal{P}(X) \to \Omega^X$ is defined as $\lambda(S : \mathcal{P}(X)).\chi_S$ where we recall that χ_S is the characteristic function of $S \subseteq X$.

Dependent product of sets

Let $\{X_i \mid i \in I\}$ be a family of sets.
Define the set $\prod_{i \in I} X_i$ to be

$$\{h \colon I \to \bigcup_{i \in I} X_i \mid \forall i\, (h(i) \in X_i)\}$$

Note that if I is a finite set, say $I = \{1, 2, \cdots, n\}$ then

$$\prod_{i \in I} X_i \cong X_1 \times X_2 \times \cdots \times X_n$$

In case where I is a finite set, if each X_i is inhabited then the cartesian product $\prod_{i \in I} X_i$ is also inhabited. But we cannot prove this for a general I.

Axiom of choice

Axiom of Choice (AC) asserts that the set $\prod_{i \in I} X_i$ is inhabited for *any* indexing set I and any family $(X_i \mid i \in I)$ of *inhabited* sets.

Warning

The axiom of choice is highly non-constructive: if a proof of a result that does not use the axiom of choice is available, it usually provides more information than a proof of the same result that does use the axiom of choice.

Logical incarnation of Axiom of Choice

Proposition

The axiom of choice is equivalent to the statement that for any sets X and Y and any formula $p(x, y)$ with free variables $x \in X$ and $y \in Y$, the sentence

$$\forall x \in X \, \exists y \in Y \, p(x, y) \Rightarrow \exists (f \colon X \to Y) \, \forall x \in X, \; p(x, f(x)) \tag{4}$$

holds.

Proof. Assume axiom of choice. Let X and Y be arbitrary sets and $p(x, y)$ any formula with free variables $x \in X$ and $y \in Y$. For each $x \in X$, define $Y_x = \{y \in Y \mid p(x, y)\}$. Note that Y_x is inhabited for each $x \in X$ by the assumption $\forall x \in X$, $\exists y \in Y$, $p(x, y)$. By the axiom of choice there exists a function $h\colon X \to \bigcup\limits_{x \in X} Y_x$ such that $h(x) \in Y_x$ for all $x \in X$. We compose the function h with the inclusion $\cup_{x \in X} Y_x \rightarrowtail Y$, which we get from the fact that $Y_x \subseteq Y$ for each $x \in X$, to obtain a function $f\colon X \to Y$. But then $p(x, f(x)) = p(x, h(x))$ is true for each $x \in X$ by definition of the sets Y_x.

Conversely, suppose that we have a family $(X_i \mid i \in I)$ of inhabited sets. Consider the cartesian product $\prod_{i\in I} X_i$. We want to show that this product is inhabited. Define

$$p(i, x) =_{\text{def}} (x \in X_i)$$

Now, we apply the sentence (4) to the sets I, $\bigcup\limits_{i\in I} X_i$ and the formula $p(i, x)$ just defined: we find a function $f\colon I \to \bigcup\limits_{i\in I} X_i$ such that $p(i, f(i))$ for all $i \in I$.

But, by definition of $p(i, x)$, we conclude that $f(i) \in X_i$ for all $i \in I$. Hence, f is a member of $\prod_{i\in I} X_i$. □

Axiom of Choice and surjections

Given a function $p\colon Y \to X$, consider the associated family $\{Y_x \mid x \in X\}$ of sets obtained by taking fibres of p at different elements of x.

Axiom of Choice and surjections

Given a function $p \colon Y \to X$, consider the associated family $\{Y_x \mid x \in X\}$ of sets obtained by taking fibres of p at different elements of x.

Lemma

A maps $p \colon Y \to X$ is surjective if and only if the fibres Y_x are inhabited for all $x \in X$.

Axiom of Choice and surjections

Given a function $p\colon Y \to X$, consider the associated family $\{Y_x \mid x \in X\}$ of sets obtained by taking fibres of p at different elements of x.

Lemma

A maps $p\colon Y \to X$ is surjective if and only if the fibres Y_x are inhabited for all $x \in X$.

Lemma

An element of $\prod_{x \in X} Y_x$ is the same thing as a section of $p\colon Y \to X$.

Axiom of Choice and surjections

Proposition

Axiom of choice is equivalent to the statement that every surjection has a section.

Proof.

Assume AC. Let $p\colon Y \to X$ be a surjection. Therefore all the fibres Y_x are inhabited. By AC, the product $\prod_{x \in X} Y_x$ is inhabited. Hence, by the last lemma above, p has a section. □

Axiom of Choice and surjections

Proposition

Axiom of choice is equivalent to the statement that every surjection has a section.

Proof.

Assume AC. Let $p\colon Y \to X$ be a surjection. Therefore all the fibres Y_x are inhabited. By AC, the product $\prod_{x\in X} Y_x$ is inhabited. Hence, by the last lemma above, p has a section. □

Axiom of Choice and surjections

Proposition

Axiom of choice is equivalent to the statement that every surjection has a section.

Proof.

Assume AC. Let $p\colon Y \to X$ be a surjection. Therefore all the fibres Y_x are inhabited. By AC, the product $\prod_{x \in X} Y_x$ is inhabited. Hence, by the last lemma above, p has a section. Conversely, suppose that every surjection has a section. Let $\{Y_x \mid x \in X\}$ be family of sets where the set Y_x is inhabited for every $x \in X$. Consider the associated function $\sqcup_{x \in X} Y_x \to X$. Note that this map is surjective by our assumption and the first lemma above. Hence, it has a section which is the same thing as an element of $\prod_{x \in X} Y_x$. Therefore AC holds. □

Suppose $f\colon A \to B$ and $g\colon Y \to X$ are functions. We say that f is (left) orthogonal to g (and, equivalently, g is right orthogonal to f) if for any two function that make the square

$$\begin{tikzcd}
A \arrow[r, "y"] \arrow[d, "f"'] & Y \arrow[d, "p"] \\
B \arrow[r, "x"'] & X
\end{tikzcd}$$

commute (i.e. $p \circ y = x \circ f$), there is a function $d\colon B \to Y$ which makes both triangles commute

$$\begin{tikzcd}
A \arrow[r, "y"] \arrow[d, "f"'] & Y \arrow[d, "p"] \\
B \arrow[r, "x"'] \arrow[ur, dotted, "d"'] & X
\end{tikzcd},$$

i.e.

$$p \circ d = x \text{ and } d \circ f = y$$

Proposition

- *Any map right orthogonal to* $\mathbf{2} \to \mathbf{1}$ *is injective.*
- *Any map right orthogonal to* $\emptyset \to \mathbf{1}$ *is surjective.*

Cantors' theorem: $A < P(A)$

Lemma

If a function $\sigma : A \to B^A$ is surjective then every function $f : B \to B$ has a fixed point.

Proof.

Because σ is a surjection, there is $a \in A$ such that $\sigma(a) = \lambda x : A \,.\, f(\sigma(x)(x))$, but then $\sigma(a)(a) = f(\sigma(a)(a))$. □

Corollary

There is no surjection $A \to P(A)$.

Let's associate to each *finite set* X a number $\mathtt{card}(X)$, called the "cardinality" of X, which measures how many (distinct) elements the set X has. We then have

- $\mathtt{card}(X + Y) = \mathtt{card}(X) + \mathtt{card}(Y)$ and
- $\mathtt{card}(X \times Y) = \mathtt{card}(X) \times \mathtt{card}(Y)$.

More generally, for any finite set I and a family of finite sets $\{X_i \mid i \in I\}$, we have

- $\mathtt{card}(\bigsqcup\limits_{i\in I} X_i) = \sum\limits_{i\in I} \mathtt{card}(X_i)$ and
- $\mathtt{card}(\prod_{i\in I} X_i) = \prod_{i\in I} \mathtt{card}(X_i)$

Questions