Contents

Chapter 1

Some Elementary Informal Set Theory

Jan. 8, 2024

Set theory is due to Georg Cantor.

"Elementary" in the title above does not apply to the body of his work, since he went into considerable technical depth in this, his new theory.

It applies however to *our* coverage as we are going to restrict ourselves to elementary topics only.

Cantor made several technical mistakes in the process of developing set theory. The next section is about the easiest to explain and most fundamental of his mistakes.

How come he made mistakes?

Actually "mistake" is too kind a term. We are talking here about contradictions. And you need just ONE to run and run away and never stop, as any logician will tell you.

If a theory can logically imply any ONE contradiction, then it can imply everything! According to such a theory EVERYTHING is a valid theorem!

How can a theory be so ill-formed?

Well, the set theory of Cantor's —unlike Euclid's Geometry 2000 years earlier— was *not* based on axioms and rigid rules of reasoning.

That's how.

Guess what: Euclidean Geometry leads to no contradictions.

DIGRESSION. "But doing mathematics by axioms AND rules of logic was not enacted seriously until after the efforts of David Hilbert in 1930s", you say.

Well, yes, and "bees cannot possibly fly". Yet, Euclid did so (logically) fly —correctly— ca. 300BC (maybe he knew Doctor Who?)

The problem with Cantor's set theory is in the conjunction of TWO omissions

1) He never delved into the question what IS a set?

2) He did not use any logical reasoning as Euclid's (he knew Geometry, presumably?)

Issue 1) is not so serious or even an issue at all IF the "nature" of the mathematical objects you are describing is determined by their axioms: FOR EXAMPLE: You don't have to define *straight line* if you give instead an axiom that says "from two distinct points passes exactly one line"! That was the approach of Euclid's.

Bertrand Russell addressed the matter of the nature of sets explicitly, which only needs logic at the level that any mathematician without training in logic uses.

He famously salvaged set theory by saying "let us *accept* that the sets we are interested in *are formed by stages; they do not just happen*".

Modern axiomatic set theory puts all its bets in issue 2 with enough axioms that the nature of sets we want to talk about jumps out of.

1.1. Russell's "Paradox"

Cantor's *naïve* (this adjective is not derogatory but is synonymous in the literature with *informal* and *non axiomatic*) set theory was plagued by *paradoxes*, the most famous of which (and the *least* "technical") being pointed out by Bertrand Russell and thus nicknamed "Russell's paradox".†

It is astounding that one of the contradictions of Cantor's set theory is so simple that you can teach it to a first year class on discrete math.

And remember that you need only ONE contradiction to destroy a theory.

†From the Greek word "paradoxo" (παράδοξο) meaning against one's belief or knowledge; a contradiction.

Cantor's set theory is the *theory of collections* (i.e., sets) of objects, as we mentioned above, terms that were neither defined *nor was it said* how they were built.†

This theory studies operations on sets, properties of sets, and aims to use set theory as the foundation *of all mathematics*. Naturally, mathematicians "do" set theory of *mathematical object collections* — not collections of birds and other beasts.

†This is not a problem *in itself*. Euclid too did not say *what* points and lines *were*; but his axioms did characterise their nature and interrelationships: For example, he started from these (among a few others) *a priori truths* (axioms): *a unique line passes through two distinct points*; also, *on any plane, a unique line l can be drawn parallel to another line k on the plane if we want l to pass through a given point A that is not on k.*

The point is:

You cannot leave out *both* what the nature of your objects is and *how* they behave/interrelate and get away with it! Euclid omitted the former but provided the latter, so all worked out.

We have learnt some elementary aspects of set theory at high school. We will learn more in this course.

1. **Variables**. Like any theory, informal or not, informal set theory —a safe variety of which we will develop here— uses *variables* just as algebra does. There is only *one type* of variable that varies over *set* and over *atomic objects* too, the latter being objects that have no set structure. For example integers. We use the names $A, B, C, \ldots$ and $a, b, c, \ldots$ for such variables, sometimes with primes (e.g., A'') or subscripts (e.g., x_{23}), or both (e.g., x'''_{22}, Y'_{42}).

2. **Notation**. *Sets given by listing.* For example, $\{1, 2\}$ is a set that contains precisely the objects 1 and 2, while

$$\{\overbrace{1}^{\text{atom}}, \overbrace{\{5, 6\}}^{\text{set}}\}$$

is a set that contains precisely the objects 1 and $\{5, 6\}$. The braces { and } are used to show the collection/set by outright listing.

So you can display small sets by listing, as in,

$$\{1, \{2, 3, 4\}, 5, \{\{6\}\}, 7, \{8, \{9\}\}\}$$

We can do better than that, in the area of notation, although a warning is fair: The "other notation" (see below) gave a lot of grief to Cantor.

Jan. 10, 2024

3. **(The "Other") Notation**. *Sets given by "defining property".* But what if we cannot (or will not) explicitly list all the members of a set?

   Then we may define what objects x get in the set/collection by having them to *pass an entrance requirement*, $P(x)$:

   > **An object x gets in the set *iff* (*if and only if*) $P(x)$ is true of said object.**

"iff" means the same thing as "is equivalent to" or "means the same thing as".

"x is in $\{x : P(x)\}$" is equivalent to saying "$P(x)$ is true".

We denote the collection/set[†] defined by the entrance condition $P(x)$ by

$$\{x : P(x)\} \tag{1}$$

but also as

$$\{x \mid P(x)\} \tag{1'}$$

reading it "the set of *all* x *such that* (this "such that" is the ":" or "|") $P(x)$ is true [or holds]"

[†]We have not yet reached Russell's result, so keeping an open mind and humouring Cantor we still allow him (us following) to call said collection a "set".

4. "$x \in A$" is the assertion that "object x is in the set A". Of course, this assertion may be true or false or "it depends", just like the assertions of algebra $2 = 2$, $3 = 2$ and $x = y$ are so (respectively).

5. $x \notin A$ is the negation of the assertion $x \in A$.

6. **Properties**

- Sets are *named* by letters of the Latin alphabet (cf. **Variables**, above).

  Naming is pervasive in mathematics as in, e.g., "let $x = 5$" in algebra.

  So we can write "let $A = \{1, 2\}$" and let "$c = \{1, \{1, 5, 6\}\}$" to give the names A and c to the two example sets above, ostensibly because we are going to discuss these sets, and refer to them often, and it is cumbersome to keep writing things like $\{1, \{1, 5, 6\}\}$.

  Names are *not permanent*;[†] they are *local* to a discussion (argument).

[†]OK, there *are* exceptions: $\emptyset$ is the permanent name for the *empty set* —the set with no elements at all— and for that set only; $\mathbb{N}$ is the permanent name of the set of all *natural numbers*.

- **Equality of sets** (repetition and permutation do not matter!)
  Two sets A and B *are equal iff they have the same members*. Thus order and multiplicity do not matter! E.g., $\{1\} = \{1,1,1\}$, $\{1,2,1\} = \{2,1,1,1,1,2\}$.

- Here is *the fundamental equivalence pertaining to definition of sets by "defining property"*:

So, if we name the set in (1) above (p.13), S, that is, if we say "let $S = \{x : P(x)\}$", then "$x \in S$ iff $P(x)$ is true"

By the way, we almost *never say* "is true" unless we want to shout out this fact.

We would simply say instead:

$$x \in S \text{ iff } P(x) \qquad (\dagger)$$

Equipped with the knowledge of the previous bullet, we see that the symbol $\{x : P(x)\}$ defines a *unique* set/collection: Well, say A and B are so defined, that is, $A = \{x : P(x)\}$ and $B = \{x : P(x)\}$. Thus

$$x \in A \overset{A=\{x:P(x)\}}{\text{ iff }} P(x) \overset{B=\{x:P(x)\}}{\text{ iff }} x \in B$$

thus

$$x \in A \text{ iff } x \in B$$

and thus $A = B$.

Let us pursue, as Russell did, the point made in the last bullet above. Take $P(x)$ to be specifically the *mathematical assertion* $x \notin x$. He then gave a *name* to

$$\{x : x \notin x\}$$

say, R. But then, by the last bullet above, in particular, the equivalence (†),

$$x \in R \text{ iff } x \notin x \tag{2}$$

If we now *believe*,[b] as *Cantor* did, that every $P(x)$ defines a *set*, *then* R *is a set.*

[b] Informal mathematics often relies on "I know so" or "I believe" or "it is 'obviously' true". Some people call "proofs" like this —i.e., "proofs" without justification(s)— "proofs by intimidation". Nowadays, with the ubiquitousness of the qualifier "fake", one could also call them "fake proofs".

What is wrong with that?

Well, if R is a set then this object has the proper *type* to be assigned (or be given as "value") into the *variable of type "math object"*, namely, x, throughout the equivalence (2) above. But this yields the contradiction

$$R \in R \text{ iff } R \notin R \tag{3}$$

This contradiction is called the Russell's Paradox.

The following is the "traditional" way to give an exposition of Russell's argument in the literature. That is, having defined

$$R = \{x : x \notin x\}$$

and thinking it to be a set, one asks:

- Is $R \in R$? An *a priori* legitimate question since R is a *set* of MATH objects and R is such an object.

  Well, if yes, then it satisfies the entrance condition $R \notin R$. *A contradiction!*

- OK, assume then the opposite of what we *assumed* in the above bullet, namely, $R \notin R$. But then R satisfies the entrance condition! So R gets in! We have $R \in R$. *A contradiction!*

So both "$R \notin R$" and "$R \in R$" are false (and hence both are true!*) *A mind boggling very very very bad situation!*

*If $R \in R$ is false then $R \notin R$ is true. But we concluded $R \notin R$ iff $R \in R$.

This and similar paradoxes motivated mathematicians to develop formal symbolic logic and look to axiomatic set theory[†] as a means to avoiding paradoxes like the above.

Other mathematicians who did not care to use mathematical logic and axiomatic theories found a way —following Russell— to do set theory *informally*, yet *safely*.

They asked *and* answered "how are sets formed?"[‡]
Read on!

[†]There are many flavours or axiomatisations of set theory, the most frequently used being the "ZF" set theory, due to Zermelo and Fraenkel.

[‡]Actually, axiomatic set theory —in particular, its axioms are— is built upon the answers this group came up with. This story is told at an advanced level in [Tou03b].

Chapter 2

Safe Set Theory

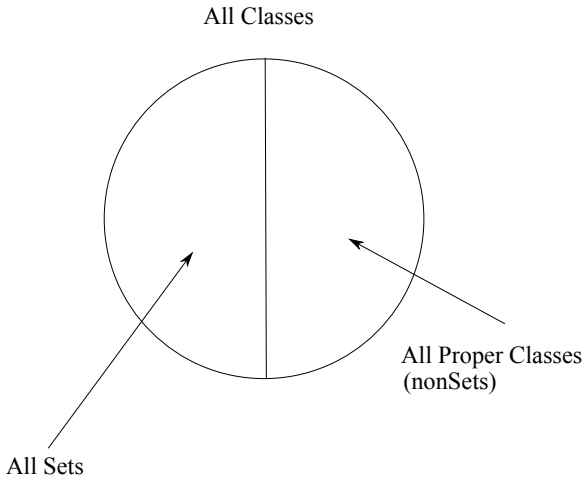
Jan. 12, 2024

So, *some* collections of sets and/or atoms are *not* —technically— sets, as the Russell Paradox taught us! How do we tell them apart?

From now on we will deal with collections that *may or may not* be sets, with a promise of learning how to create *sets* if we want to!

The modern literature uses the terminology "**class**" for *any* such collection of sets and/or atoms (and uses the term "collection" non technically and sparsely).

The above is captured by the following picture:



So *some* classes are *proper* (*NON sets*) and some are not (i.e., ARE sets).

So *every set is a class* but *NOT the other way around!*

2.0.1 Definition. (Classes and sets)
From now on we call *all* collections **classes**.

Definitions by defining property like "Let $\mathbb{A} = \{x : P(x)\}$", where x is a set/atom-type variable, **always** define a **class**, but as we saw, sometimes —e.g., when "$P(x)$" is specifically "$x \notin x$"— that class is *not* a set (Section 1.1).

Classes that are not sets are called **proper classes**.

The "property" $x \notin x$ is not "cursed"! Infinitely many properties define **PROPER** classes. As we will shortly see, the property "$x = x$" defines a proper class too.

We will normally use what is known as "blackboard bold" notation and capital latin letters to denote classes by names such as $\mathbb{A}, \mathbb{B}, \mathbb{X}$. If we determine that some class $\mathbb{A}$ *is* a set, we would rather write it as A, but we make an *exception* for the following **sets**:

The set of natural numbers, $\mathbb{N}$ (also denoted by ω), integers $\mathbb{Z}$, rationals $\mathbb{Q}$, reals $\mathbb{R}$ and *complex numbers* $\mathbb{C}$. □

2.0.2 Example. By the Definition just given, if R is the Russel (proper) class, then the configuration

$$\{R\}$$

is not allowed —it is meaningless.

> Because ALL classes are collections of **atoms and sets** only. We *never said that it is OK, and will* NEVER *allow*, proper classes as *members* of classes!

Of course Cantor would not care and allow $\{R\}$ and even this

$$\{\{\{R\}\}, R\}$$

because *in his set theory ALL collections were "sets" or "classes" or "aggregates" or ...* (just give me a Dictionary!) □

In forming the class $\{x : P(x)\}$ for any property $P(x)$ we say that we apply *comprehension*. It was the Frege/Cantor belief (explicitly or implicitly) that comprehension was *safe* —i.e., they believed that $\{x : P(x)\}$ always was a set. We saw that Russell proved that it was not.

2.1. The "real sets" —Introduction to Stages

So, how can we tell, or indeed *guarantee*, that a certain *class* is a *set*?

Russell proposed this "recovery" from his Paradox:

Make sure that sets are built by stages, where at stage 0 all atoms are available.

We may then collect atoms to form all sorts of "first level" *sets*. We may proceed to collect any mix of atoms and first-level sets to build new collections —second-level sets— *and so on*.

Much of what set theory does is attempting to remove any ambiguity from this "and so on". See below, **Principles** 0–2.

Thus, at the beginning we have all the level-0, or type-0, objects available to us. For example, *atoms* such as $1, 2, 13, \sqrt{2}$ are available.

At the next level we can include any number of such atoms (from none at all, to all) to **build a set**, that is, a new mathematical object.

Allowing the usual notation, i.e., **listing** of what is included within braces, we may cite a few examples of level-1 **sets**:

L1-1. $\{1\}$.

L1-2. $\{1, 1\}$.

L1-3. $\{1, \sqrt{2}\}$.

L1-4. $\{\sqrt{2}, 1\}$.

L1-5. $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

We already can identify a level-2 object, using what (we already know) *is* available:

L2-1. $\{\{\sqrt{2}, 1\}, 42\}$.

Note how the level of nesting of { }-brackets matches the level or stage of the formation of these objects!

2.1.1 Definition. (Class and set *equality* —again) This definition applies to any classes, hence, in particular, to any *sets* as well.

Two classes $\mathbb{A}$ and $\mathbb{B}$ are *equal* —written $\mathbb{A} = \mathbb{B}$— means

$$x \in \mathbb{A} \text{ iff } x \in \mathbb{B} \tag{1}$$

That is, an object is in $\mathbb{A}$ IF it is also in $\mathbb{B}$.

And, an object is in $\mathbb{B}$ IF it is also in $\mathbb{A}$.

$\mathbb{A}$ is a *subclass* of $\mathbb{B}$ —written $\mathbb{A} \subseteq \mathbb{B}$— means that every element of the first (left) class occurs also in the second, or

$$\text{If } x \in \mathbb{A}, \text{ then } x \in \mathbb{B} \tag{2}$$

If $\mathbb{A}$ is a *set*, then we say it is a *subset* of $\mathbb{B}$.

If we have $\mathbb{A} \subseteq \mathbb{B}$ but $\mathbb{A} \neq \mathbb{B}$, then we write $\mathbb{A} \subsetneqq \mathbb{B}$ (some of the literature uses $\mathbb{A} \subsetneq \mathbb{B}$ or even $\mathbb{A} \subset \mathbb{B}$ instead) and say that $\mathbb{A}$ is a *proper subclass* of $\mathbb{B}$.

Caution. In the terminology "*proper subclass*" the "proper" refers to the fact that $\mathbb{A}$ is not all of $\mathbb{B}$. It does *NOT* say that $\mathbb{A}$ is not a set! It *may* be a set and then we say that it is a "*proper subset*" of $\mathbb{B}$ □

If n is an integer-valued variable, then what do you understand by the statement "$2n$ is even"?

The normal understanding is that "no matter what the value of n is, $2n$ is even", or "for all values of n, $2n$ is even".

When we get into our logic topic in the course we will see that we *can* write "for all values of n, $2n$ is even" with less English as "$(\forall n)(2n$ is even)". So "$(\forall n)$" says "for all (values of) n".

Mathematicians often prefer to have statements like "$2n$ is even" with the "for all" *only implied.*[†] You can write a whole math book without writing $\forall$ even once, and without overdoing the English.

Thus in (1) and (2) above the "for all x" is implied.
For example, this is the intend in $x \in \mathbb{A} \to x \in \mathbb{B}$ and $x \in \mathbb{A} \equiv x \in \mathbb{B}$.

But in "Let $x \in \mathbb{A}$" we speak of an unspecified FIXED value of x.

[†]An exception occurs in Induction that we will study later, where you *fix* an n (but keep it as a variable of an unspecified fixed value; not as 5 or 42) and assume the "induction hypothesis" $P(n)$. But do not worry about this now!

2.1.2 Remark. Since "iff" or "$\equiv$" between two statements S_1 and S_2 means that we have *both* directions —boxed statement in 2.1.1,

$$\text{If } S_1, \text{ then } S_2$$

and

$$\text{If } S_2, \text{ then } S_1$$

we have that "$\mathbb{A} = \mathbb{B}$" is the same as (*equivalent to*) "$\mathbb{A} \subseteq \mathbb{B}$ *and* $\mathbb{B} \subseteq \mathbb{A}$" (2.1.1).

This is because (1) in 2.1.1 means $x \in \mathbb{A} \to x \in \mathbb{B}$ *AND* $x \in \mathbb{B} \to x \in \mathbb{A}$. □

2.1.3 Example. In the context of the "$\mathbb{A} = \{x : P(x)\}$" notation we should remark that notation-by-listing can be simulated by notation-by-defining-property: For example, $\{a\} = \{x : x = a\}$ —here "$P(x)$" is $x = a$.

Also $\{A, B\} = \{x : x = A \text{ or } x = B\}$. Let us verify the latter: Say $x \in$ lhs.[†] Then $x = A$ or $x = B$. But then the entrance requirement of the rhs[‡] is met, so $x \in$ rhs.

Conversely, say $x \in$ rhs. Then the entrance requirement is met so we have (at least) one of $x = A$ or $x = B$ ("true" implied).
Trivially, in the first case $x \in$ lhs and ditto for the second case. □

[†]Left Hand Side.
[‡]Right Hand Side.

Jan. 15, 2024

We now postulate the principles of formation of sets!

Principle 0.
Sets are formed by STAGES. At stage 0 we have the *presence* of ALL atoms. *They are given outright, they are not built.*

> At *any* stage Σ we *are allowed to* build a *set*, collecting together other *mathematical objects* (sets or atoms) *provided* (*iff*) *ALL these (mathematical) objects that we put into our set were ALL available at stages BEFORE* Σ.

Principle 1. *EVERY set is built at SOME stage. Thus, a set does not just happen!*

Principle 2. If Σ is a stage of set construction, then *there IS* a stage Φ *after* it.

We can write this as "$\Sigma < \Phi$".

Principle 2 makes clear that we have *infinitely many* stages of set formation in our toolbox.

"Clear"?

Can you argue that informally? (**Exercise**! *Hint.* Combine Property 2 statement with a "what if": *What if there are only finitely many stages?* and go for a contradiction from the what if. Use the "obvious" properties of $<$ *between stages* that we postulate below.)

Incidentally the property of a stage being "before" another is exactly like "$<$" on the integers:

1. For any two integers n, m the statement "$n = m$ or $n < m$ or $m < n$" is true.
2. We cannot have $n < n$, for any n (this is the "irreflexivity" of "$<$").
3. If we have $n < m$ and $m < r$, then we also have $n < r$ (this is the "transitivity" of "$<$").

For stages,

Using "$<$" as short for "lhs comes *before* rhs", then

1′. For any two stages Σ and Σ' the statement "$\Sigma = \Sigma'$ or $\Sigma < \Sigma'$ or $\Sigma' < \Sigma$" is true.

2′. We cannot have that Σ is before (or after) Σ, for any Σ.

3′. If we have $\Sigma < \Sigma'$ and $\Sigma' < \Sigma''$, then $\Sigma < \Sigma''$.

2.1.4 Remark. If some set is definable ("buildable") at some stage Σ, then it is also definable at any later stage as well, as **Principle** 0 makes clear.

The informal set-formation-by-stages Principle will guide us to build, safely, all the sets we may need in order to do mathematics.

□

2.2. What caused Russell's paradox

How would the set-building-by-stages doctrine avoid Russell's paradox?

Recall that *à la Cantor* we get a paradox (*contradiction*, actually) because we *insisted to believe* that **ALL classes are sets**, that is, following Cantor we "believed" (we just pretended!) —for a short moment— that Russell's "R" was a *set*.

Principles 0–2 allow us to know *a priori* that R is a proper class. **BEFORE** any contradiction occurs!

How so?

OK, **FIRST** let us ask and explore: is $x \in x$ **true** or **false**? Is there *any* mathematical object x —say, A— for which it *is* true?

$$A \in A? \tag{1}$$

1. Well, for atom A, (1) is false since *atoms have no set structure*, that is, they do NOT contain ANY objects: An atom A *cannot contain anything*, in particular it cannot contain A.

2. What if A is a **set** and $A \in A$? Then in order to build A, the *set* on the rhs, we have to wait until *after* its member, A —the set on the lhs— is built (Principle 0). So, we need (the left) A to be built *BEFORE* (the right) A in (1).

Absurd!

So (1) is **false**. A being arbitrary, we have just demonstrated that

$$x \in x \text{ is false (for all } x \text{ that are sets or atoms).}$$

thus $x \notin x$ is true (*for all* x) —just like $x = x$ is— therefore R of Section 1.1 is equal to $\mathbb{U}$ —they both have as "entrance condition" a property that is **always true**: We could write $R = \mathbb{U} = \{x : \mathbf{t}\}$.

By $\mathbb{U}$ we denote the universe of *all sets and atoms*.

$$R = \mathbb{U} = \{x : x = x\}$$

So?

SECOND,

So here is why we know that $\mathbb{U}$ —that is, R— is *not* a set. Well, if it is, then

- $\mathbb{U} \in \mathbb{U}$ since the rhs contains *all sets* and we *assumed* the lhs to be a *set*.
- but we just saw that the above is false if $\mathbb{U}$ is a *set*!

So $\mathbb{U}$, aka R, is a *proper* class. Thus, the fact that R is not a set is neither a surprise, nor paradoxical. It is just a *proper* class as we just have recognised WITHOUT REPEATING Russell's ARGUMENT.

BTW,

A class $\mathbb{A}$ is proper iff we have *NO stage left to build it* (Principles 0 and 1).

Intuitively then if we ran out of stages building $\mathbb{A}$ it means that *there are far too many elements in* $\mathbb{A}$ —this class is "enormous", as indeed $\mathbb{U} = \Big\{x : x = x\Big\}$ is.

Often the informal (and sloppy) literature on sets will blame "size" for a class failing to be a set. That is dangerous. Lack of set status must be connected with *lack of a stage* at which to build said class as a set.

Incidentally not all "LARGE" classes contain "everything". We will see later that if we remove ALL atoms from $\mathbb{U}$, then what remains is a proper class too. So is $\mathbb{S} = \{\{x\} : x \in \mathbb{U}\}$: The class of *all 1-element sets*. It is much smaller than $\mathbb{U}$: No 2-element sets, no 3-element sets, no infinite set objects in $\mathbb{S}$ either!

2.3. Some useful sets

2.3.1 Example. (Pair) By Principles 0, 1, if A and B are sets or atoms, then let A be available at stage Σ and B at stage Σ'.

There are just two cases (just two? Why?)

By Principle 2 take a new $\Sigma'' > \Sigma'$ in each case below.

Case **1.** $\Sigma < \Sigma'$. Then also $\Sigma < \Sigma''$ by *transitivity*. So both A and B are built or available *BEFORE* Σ'' and we can build (Princ. 0!) $\{A, B\}$ as a *SET* at stage Σ''.

Case **2.** $\Sigma = \Sigma'$. As before, by Principle 2, we take $\Sigma'' > \Sigma'$.

But then also $\Sigma < \Sigma''$ (Why?)

So both A and B are built or available *BEFORE* stage Σ'' and we can build (Princ. 0!) $\{A, B\}$ as a *SET* at stage Σ''. □

Pause. We call $\{A, B\}$ the "(unordered) *Pair*"

Why "unordered"? See 2.1.1.◀

We have just proved a theorem above:

2.3.2 Theorem. *If A, B are sets or atoms, then $\{A, B\}$ is a set.*

2.3.3 Exercise. Without referring to stages in your proof, prove that if A is a set or atom, then $\{A\}$ is a set. □

Jan. 17, 2024

2.3.4 Remark. A very short digression into Boolean Logic — for now. It will be convenient to use *truth tables* to handle many simple situations that we will encounter where "logical connectives" such as "*not*", "*and*", "*or*", "*implies*" and "*is equivalent*" enter into our arguments.

We will put on record here how to *compute* things such as the **true/false value** —called "*truth-value*"— of "S_1 and S_2", "S_1 or S_2", etc., where S_1 and S_2 stand for two arbitrary statements of mathematics.

In the process we will introduce the *mathematical symbols* for "and", "implies", etc.

The *symbol translation table* from English to symbol, and back, is:

NOT	$\neg$
AND	$\wedge$
OR	$\vee$
IMPLIES (IF...,THEN)	$\rightarrow$
IS EQUIVALENT	$\equiv$

The truth table below has a simple reading. For *all possible* truth values —**true**/**false**, in short **t**/**f**— of the "simpler" statements S_1 and S_2 we indicate the *computed truth value* of the compound (or "more complex)" statement that we obtain when we *apply* one or the other **Boolean connective** —I also call this "glue" in my logic course :)— of the previous table to S_1 and S_2.

Table 2.1: Truth Tables

S_1	S_2	$\neg S_1$	$S_1 \wedge S_2$	$S_1 \vee S_2$	$S_1 \to S_2$	$S_1 \equiv S_2$	$S_2 \to S_1$
f	**f**	**t**	**f**	**f**	**t**	**t**	**t**
f	**t**	**t**	**f**	**t**	**t**	**f**	**f**
t	**f**	**f**	**f**	**t**	**f**	**f**	**t**
t	**t**	**f**	**t**	**t**	**t**	**t**	**t**

Comment. All the computations of truth values *satisfy our intuition*, with the *possible* —but not necessary— exception for "$\rightarrow$":

Indeed, $\neg$ flips the truth value as it should, $\wedge$ is eminently consistent with common sense, $\vee$ is the "inclusive or" —"**this is true or the other is true OR both**"— of the mathematician, and $\equiv$ is just equality on the set $\{\mathbf{f}, \mathbf{t}\}$, as it should be: **we have $S_1 \equiv S_2$ true EXACTLY IF both S_i are t or both are f.**

The "problem" with $\rightarrow$ is that there is no **NECESSARILY** *causality* from left to right. The only "obvious" entry seems to be for $\mathbf{t} \rightarrow \mathbf{f}$. The outcome should be false for a "bad implication"[†] and so it is.

But look at it this way:

- Implication is supposed to preserve truth in proofs.
  But it does do just that! Just look at $\rightarrow$ truth column!
- This version of $\rightarrow$ goes way back to Aristotle. It is the version used by the vast majority of practising mathematicians and is nicknamed "material implication" or "classical implication".

[†]A bad implication has a true premise but a false conclusion. A correct implication ought to preserve truth!

Practical considerations. Thus

1. if you want to demonstrate that $S_1 \vee S_2$ is true, for any component statements S_1, S_2, then show that *at least one* of the S_1 and S_2 is true.

2. If you want to demonstrate that $S_1 \wedge S_2$ is true, then show that *both* of the S_1 and S_2 are true.

   Note, incidentally, the if we *know* that $S_1 \wedge S_2$ is true, then the truth table *guarantees* that each of S_1 and S_2 *must* be true.

3.

   If now you want to show the implication $S_1 \rightarrow S_2$ is true, **then the ONLY real work is required towards showing that *if we assume* S_1 is true, then S_2 is true too**.

   If S_1 is known to be false, then no work is required to prove the implication because of the first two lines of the truth table!!

4. If you want to show $S_1 \equiv S_2$, then —since the last three columns show that this is *computed* with *the same result* as $\Big(S_1 \rightarrow S_2\Big) \wedge \Big(S_2 \rightarrow S_1\Big)$— it follows that you just have to *compute* and "show" that **each** of the two implications $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_1$ is true.

   Priorities and Bracketing. Priority order is

$$\neg, \wedge, \vee, \rightarrow, \equiv$$

How do I compute $2 + 3 \times 4$?

Analogously, $A \vee B \wedge C$ says $A \vee (B \wedge C)$, $\neg A \vee B$ says $(\neg A) \vee B$, $A \equiv B \equiv C$ says $A \equiv (B \equiv C)$, $A \to B \to C$ says $A \to (B \to C)$, $A \vee B \vee C$ says $A \vee (B \vee C)$ (*right associativity*).

An important variant of $\to$ and $\equiv$

Pay attention to this point since almost everybody gets it wrong! In the literature and in the interest of creating a usable shorthand many practitioners of mathematical writing use sloppy notation

$$S_1 \to S_2 \to S_3 \tag{1}$$

attempting to convey the meaning

$$(S_1 \to S_2) \wedge (S_2 \to S_3) \tag{2}$$

Alas, (2) is not the same as (1)! But what about writing $a < b < c$ for $a < b \wedge b < c$? *That is wrong too!*

Back to $\to$-chains like (1) vs. chains like (2):

Take S_1 to be $\mathbf{t}$ (true), S_2 to be $\mathbf{f}$ and S_3 to be $\mathbf{t}$. Then (1) is true because in a chain using the same Boolean connective *we put brackets from right to left*: (1) says $S_1 \to (S_2 \to S_3)$ and evaluates to $\mathbf{t}$, while (2) evaluates clearly to false ($\mathbf{f}$) since $S_1 \to S_2 = \mathbf{f}$ and $S_2 \to S_3 = \mathbf{t}$.

So we need a special symbol to denote (2) "*economically*". We need a *conjunctional implies*! Most people use "$\Longrightarrow$" for that:

$$S_1 \Longrightarrow S_2 \Longrightarrow S_3 \tag{3}$$

that means, by **definition**, (2) above.

Similarly,

$$S_1 \equiv S_2 \equiv S_3 \tag{4}$$

is **NOT** conjunctional. It is **not** two equivalences —two statements— connected by an *implied* "$\wedge$", rather it says

$$S_1 \equiv (S_2 \equiv S_3)$$

ONE formula, ONE statement.

Now if $S_1 = \mathbf{f}$, $S_2 = \mathbf{f}$ and $S_3 = \mathbf{t}$, then (4) evaluates as $\mathbf{t}$ but the conjunctional version

$$(S_1 \equiv S_2) \wedge (S_2 \equiv S_3) \tag{5}$$

evaluates as $\mathbf{f}$ since the second side of $\wedge$ is $\mathbf{f}$.

So how do we denote (5) correctly without repeating the consecutive S_2's and omitting the implied "$\wedge$"? This way:

$$S_1 \Longleftrightarrow S_2 \Longleftrightarrow S_3 \tag{4}$$

By definition, "$\Longleftrightarrow$" —just like "iff"— is conjunctional: It applies to two statements —S_i and S_{i+1}— only and implies an $\wedge$ before the adjoining next similar equivalence. □

Jan. 19, 2024

2.3.5 Theorem. (The subclass theorem) *Let $\mathbb{A} \subseteq B$ (B a set). Then $\mathbb{A}$ is a set.*

Proof. Well, B being a set it is built at some state Σ (Principle 1).

By Principle 0, ALL members of B are *available or built* before *stage* Σ.

But by $\mathbb{A} \subseteq B$, *ALL the members of* $\mathbb{A}$ *are among those of* B.

So all members of $\mathbb{A}$ are built/available BEFORE stage Σ.

Hey! By Principle 0 we can build $\mathbb{A}$ at stage Σ as a set. □

Some corollaries are very useful:

2.3.6 Corollary. (Important!) *If B is built at stage Σ then EACH of its subclasses can be built at stage Σ as well.*

2.3.7 Corollary. (Modified comprehension I) *If for all x we have*

$$P(x) \to x \in A \tag{1}$$

for some SET A, then it is SAFE *to build*

$$\mathbb{B} = \{x : P(x)\} \tag{†}$$

as a SET. *No funny business with the condition "$P(x)$".*

Proof. I will show that $\mathbb{B} \subseteq A$, that is,

$$x \in \mathbb{B} \to x \in A \tag{2}$$

Let's do the above in two implication steps using the conjunctional implication "$\Rightarrow$":

$$x \in \mathbb{B} \overset{by\ (\dagger)}{\Rightarrow} P(x) \overset{by\ (1)}{\Rightarrow} x \in A \tag{3}$$

(3) proves (2). □

2.3.8 Corollary. (Modified comprehension II) *If A is a set, then so is $\mathbb{B} = \{x : x \in A \wedge P(x)\}$ for any property $P(x)$.*

Proof. The defining property here is "$(x \in A)^{\dagger} \wedge P(x)$". This implies $x \in A$ —by 2 in 2.3.4— that is, we have

$$x \in A \wedge P(x) \to x \in A$$

Now invoke 2.3.7. □

[†]Brackets not needed; inserted for extra clarity.

2.3.9 Remark. (*The* empty set) The class $\mathbb{E} = \{x : x \neq x\}$ has no members at all; it is empty. Why? Because

$$x \in \mathbb{E} \equiv x \neq x$$

but the condition $x \neq x$ is *always false*, therefore *so is the statement*

$$x \in \mathbb{E} \tag{1}$$

We do not collect anything into $\mathbb{E}$. Is the class $\mathbb{E}$ a set?

Well, take $A = \{1\}$. This is a set as the atom 1 is given at stage 0, and thus we can construct the *set* A at stage 1.

Note that, by (1) and 3 in 2.3.4 we have that the implication below

$$\overbrace{x \in \mathbb{E}}^{\mathbf{f}} \underbrace{\rightarrow}_{\mathbf{t}} x \in \{1\}$$

is true (for all x). That is, $\mathbb{E} \subseteq \{1\}$.

By 2.3.5, $\mathbb{E}$ *is a set.*

But is it *unique* so we can justify the use of the definite article "the"? **Yes**. The specification of *an empty set is a class with no members*. So if D is another empty set, then we will *also* have $x \in D$ always *false*. But then

$$\overbrace{x \in \mathbb{E}}^{\mathbf{f}} \underbrace{\equiv}_{\mathbf{t}} \overbrace{x \in D}^{\mathbf{f}}$$

and we have $\mathbb{E} = D$ by 2.1.1.

The unique *empty set is denoted by the* symbol $\emptyset$ *in the literature*.

Never ever use "{}" for the empty set. This incorrect notation is used —as everything else sloppy and wrong— in fake math news! □ 

2.4. Operations on classes and sets

The reader probably has seen before (perhaps in calculus) the operations on sets denoted by $\cap, \cup, -$ and others. We will look into them in this section.

2.4.1 Definition. (Union of two classes) We define for any classes $\mathbb{A}$ and $\mathbb{B}$

$$\mathbb{A} \cup \mathbb{B} \stackrel{Def}{=} \Big\{x : x \in \mathbb{A} \vee x \in \mathbb{B}\Big\}$$

We call the operator $\cup$ *union* and the result $\mathbb{A} \cup \mathbb{B}$ the union of $\mathbb{A}$ and $\mathbb{B}$.

It is meaningless to have $\cup$ operate on atoms. □

2.4.2 Theorem. *For any sets A and B, $A \cup B$ is a set.*

Proof. By assumption —"sets", we assumed!— say, A is built at stage Σ while B is built at stage Σ'.

As in the proof in Example 2.3.1, Principle 2 guarantees a stage Σ'' such that

$$\Sigma < \Sigma'' \tag{1}$$

and

$$\Sigma' < \Sigma'' \tag{2}$$

Now let us pick any item $x \in A \cup B$:

I have two (not necessarily mutually exclusive) cases* (by 2.4.1):

- $x \in A$. Then x was available or built BEFORE Σ'' by (1).†
- $x \in B$. Then x was available or built BEFORE Σ'' by (2).‡

Thus ALL x in $A \cup B$ can form a *set* at stage Σ''. □

*The "or both" case reduces to case "$x \in A$", trivially (x is in both, then it ios in A).
†Because $x \in A$ is available BEFORE Σ. Now use (1) and transitivity of $<$.
‡Because $x \in B$ is available BEFORE Σ'. Now use (2) and transitivity of $<$.

Jan. 22, 2024

2.4.3 Definition. (Intersection of two classes) We define for any classes $\mathbb{A}$ and $\mathbb{B}$

$$\mathbb{A} \cap \mathbb{B} \overset{Def}{=} \Big\{ x : x \in \mathbb{A} \wedge x \in \mathbb{B} \Big\} \tag{1}$$

We call the operator $\cap$ *intersection* and the result $\mathbb{A} \cap \mathbb{B}$ the intersection of $\mathbb{A}$ and $\mathbb{B}$.

If $\mathbb{A} \cap \mathbb{B} = \emptyset$ —which happens precisely when the two classes have no common elements— we call the classes *disjoint*.

Taking liberties with notation (of definition by defining property) we may write instead of (1) either

$$\mathbb{A} \cap \mathbb{B} \overset{Def}{=} \Big\{ x \in \mathbb{A} : x \in \mathbb{B} \Big\} \tag{1'}$$

or

$$\mathbb{A} \cap \mathbb{B} \overset{Def}{=} \Big\{ x \in \mathbb{B} : x \in \mathbb{A} \Big\} \tag{1''}$$

It is meaningless to have $\cap$ operate on atoms.[†] □

We have the easy theorem below:

[†]The definition expects $\cap$ to *operate on classes*. As we know, atoms (by definition) *have no set/class structure* thus no class and no set is an atom.

2.4.4 Theorem. *If B is a set, as its notation suggests, then $\mathbb{A} \cap B$ is a set.*

Proof. I will prove $\mathbb{A} \cap B \subseteq B$ which will rest the case by 2.3.5. So, I want

$$x \in \mathbb{A} \cap B \to x \in B$$

To this end, let then $x \in \mathbb{A} \cap B$ (cf. 3 in 2.3.4).

This says that $x \in \mathbb{A} \wedge x \in B$ is true. Hey! So $x \in B$ is true. □

2.4.5 Corollary. *For sets A and B, $A \cap B$ is a set.*

2.4.6 Definition. (Difference of two classes) We define for any classes $\mathbb{A}$ and $\mathbb{B}$

$$\mathbb{A} - \mathbb{B} \overset{Def}{=} \Big\{x : x \in \mathbb{A} \wedge x \notin \mathbb{B}\Big\} \tag{1}$$

We call the operator "$-$" *difference* and the result $\mathbb{A} - \mathbb{B}$ the difference of $\mathbb{A}$ and $\mathbb{B}$, in that order.

It is meaningless to have "$-$" operate on atoms. □

Notation. As was the case for $\cap$ (Definition 2.4.3) for "$-$" too we have a shorter alternative notation to (1) above:

$$\mathbb{A} - \mathbb{B} \overset{Def}{=} \Big\{x \in \mathbb{A} : x \notin \mathbb{B}\Big\}$$

2.4.7 Theorem. *For any* set A *and class* $\mathbb{B}$, $A - \mathbb{B}$ *is a set.*

Proof. The reader is asked to verify that $A - \mathbb{B} \subseteq A$. We are done by 2.3.5. □

2.4.8 Exercise. Prove that $\{\mathbb{Z}\}$ is a set, where $\mathbb{Z}$ is the set of integers $\{\ldots, -1, 0, 1, \ldots\}$. □

2.4.9 Exercise. Demonstrate —using Definition 2.4.3— that for any $\mathbb{A}$ and $\mathbb{B}$ we have $\mathbb{A} \cap \mathbb{B} = \mathbb{B} \cap \mathbb{A}$.

Hint. You can do this by doing

$$x \in \mathbb{A} \cap \mathbb{B} \to x \in \mathbb{B} \cap \mathbb{A} \text{ (for all } x)$$

This is *normally* done by fixing an x and going "Let $x \in \mathbb{A} \cap \mathbb{B}$. Then BLA BLA BLA, therefore $x \in \mathbb{B} \cap \mathbb{A}$", and then repeating the argument backwards: "Let $x \in \mathbb{B} \cap \mathbb{A}$. ETC."

OR you could note the definition for $\mathbb{A} \cap \mathbb{B}$, that is, $= \left\{x : x \in \mathbb{A} \wedge x \in \mathbb{B}\right\}$ AND the definition for $\mathbb{B} \cap \mathbb{A}$ and prove by truth tables that the defining properties of the two are EQUIVALENT (easy!!!) □

2.4.10 Exercise. Demonstrate —using Definition 2.4.1— that for any $\mathbb{A}$ and $\mathbb{B}$ we have $\mathbb{A} \cup \mathbb{B} = \mathbb{B} \cup \mathbb{A}$. □

2.4.11 Exercise. By picking two particular very small sets A and B show that $A - B = B - A$ is not true for all sets A and B.

Is it true of all classes? □

2.5. The powerset

2.5.1 Definition. For any set A the symbol $\mathscr{P}(A)$ —pronounced the *powerset* of A— is defined to be the class

$$\mathscr{P}(A) \stackrel{Def}{=} \Big\{ x : x \subseteq A \Big\}$$

Thus we collect *all* the subsets x of A to form $\mathscr{P}(A)$.

The literature most frequently uses the symbol 2^A in place for $\mathscr{P}(A)$. □

(1) The term "power*set*" is slightly premature, but it is apt. Under the conditions of the definition —A a set— 2^A is a *set* as we prove immediately below.

(2) We said "*all* the sub*sets* x of A" in the definition. This is correct. As we know from 2.3.5, if $\mathbb{X} \subseteq Y$ and Y is a set, then so is $\mathbb{X}$.

2.5.2 Theorem. *For any set A, its powerset $\mathscr{P}(A)$ is a set.*

Proof. Let A be built at stage Σ.

By 2.3.6, if $x \subseteq A$ then x can be build at stage Σ. Well, let us by Princ. 2, pick a stage Σ' *after* Σ: That is, $\Sigma < \Sigma'$.

Hence each $x \subseteq A$ can be built before Σ'. Then we can collect all these x in a *SET*!

That *set* is $\{x : x \subseteq A\} = 2^A$. □

2.5.3 Example. Let $A = \{1, 2, 3\}$. Then

$$\mathscr{P}(A) = \Big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{3, 2\}, \{1, 2, 3\}\Big\}$$

Thus the powerset of A has 8 elements.

We will later see that if A has n elements, for any $n \geq 0$, then 2^A has 2^n elements. This observation is at the root of the notation "2^A". □

2.5.4 Remark. For any set A it is trivial (verify!) that we have $\emptyset \subseteq A$ and $A \subseteq A$. Thus, for any A, $\{\emptyset, A\} \subseteq 2^A$. □

Let us generalise unions and intersections next. First a definition:

2.5.5 Definition. (Families of sets) A class $\mathbb{F}$ is called a *family of sets* iff *it contains NO atoms*. The letter $\mathbb{F}$ is here used generically —$\mathbb{F}$ for "family"— and a family may be given any name, usually capital (blackboard bold if we do not know that it is a set). □

2.5.6 Example. Thus, $\emptyset$ is a family of sets; the empty family.

So are $\{\{2\}, \{2, \{3\}\}\}$ and $\mathbb{V}$, the latter given by

$$\mathbb{V} \overset{Def}{=} \Big\{ x : x \text{ is a set} \Big\}$$

BTW, as $\mathbb{V}$ contains all sets (but no atoms!) it is a proper class!

Why? Well, if it is a set, then it is one of the x-values that we are collecting, thus $\mathbb{V} \in \mathbb{V}$. But we saw that this statement is false for sets!

Here are some classes that are *NOT* families: $\{1\}$, $\{2, \{\{2\}\}\}$ and $\mathbb{U}$, the latter being the universe of all objects —sets *and* atoms— and equals Russell's "R" as we saw in Section 2.2. These all are disqualified as they contain atoms. □

2.5.7 Definition. (Intersection and union of families) Let $\mathbb{F}$ be a family of sets. Then

(i) the symbol $\bigcap\mathbb{F}$ denotes the class that contains *all the objects* x that *are FOUND in* ALL $A\in\mathbb{F}$.

In symbols the definition reads:

$$\bigcap\mathbb{F}\overset{Def}{=}\Big\{x:\text{for all } A, A\in\mathbb{F}\to x\in A\Big\} \tag{1}$$

(ii) the symbol $\bigcup\mathbb{F}$ denotes the class that contains *all the objects* that *are found among the various* $A\in\mathbb{F}$. That is, imagine that the members of *each* $A\in\mathbb{F}$ are "emptied" into a single —originally empty— container $\{\ldots\}$. The class we get this way is what we denote by $\bigcup\mathbb{F}$.

In symbols the definition reads (and I think it is clearer):

$$\bigcup\mathbb{F}\overset{Def}{=}\Big\{x:\text{for some } A, A\in\mathbb{F}\wedge x\in A\Big\} \tag{2}$$

$$\text{So include } x \textbf{ iff } x\in \overset{\overset{any}{\downarrow}}{A}\in\mathbb{F}$$

So ALL $x\in A\in\mathbb{F}$ ARE collected!

□

2.5.8 Example. Let $\mathbb{F} = \{\{1\}, \{1, \{2\}\}\}$. Then emptying all the contents of the members of $\mathbb{F}$ into some (originally) empty container we get

$$\{1, 1, \{2\}\} \tag{3}$$

This is $\bigcup \mathbb{F}$.

Would we get the same answer from the mathematical definition (2)? Of course:

1 *is* in some member of $\mathbb{F}$, indeed in both of the members $\{1\}$ and $\{1, \{2\}\}$, and in order to emphasise this I wrote two copies of 1 —it is emptied/contributed twice. Then $\{2\}$ is the member that only $\{1, \{2\}\}$ of $\mathbb{F}$ contributes.

We do not see any other members in the two set-members —$\{1\}$ and $\{1, \{2\}\}$— of $\mathbb{F}$. So, all done!

What is $\bigcap \mathbb{F}$? Well, only 1 is the only one common between the two sets —$\{1\}$ and $\{1, \{2\}\}$— that are in $\mathbb{F}$. So, $\bigcap \mathbb{F} = \{1\}$. □

2.5.9 Exercise.

The below four operations were defined independently of each other. Let us compare them:

1. Prove that $\bigcup \Big\{ A, B \Big\} = A \cup B$.
2. Prove that $\bigcap \Big\{ A, B \Big\} = A \cap B$.

Hint. In each of part 1. and 2. show that lhs $\subseteq$ rhs and rhs $\subseteq$ lhs. For that analyse membership, i.e., "assume $x \in$ lhs and prove $x \in$ rhs", and conversely (cf. 2.1.1 and 2.1.2.) □

2.5.10 Theorem. *If the class $\mathbb{F} \neq \emptyset$ is a family of sets, then $\bigcap \mathbb{F}$ is a set.*

Proof. By assumption there is some set in $\mathbb{F}$. Fix *one* such and call it D.

First note that

$$x \in \bigcap \mathbb{F} \to x \in D \qquad (*)$$

Why? Because (1) of Definition 2.5.7 (1) says that

$$x \in \bigcap \mathbb{F} \equiv \text{for all } A \in \mathbb{F} \text{ we have } x \in A$$

Well, D *is* one of those "A" sets in $\mathbb{F}$, so if $x \in \bigcap \mathbb{F}$ then $x \in D$. We established $(*)$ and thus we established

$$\bigcap \mathbb{F} \subseteq D$$

by 2.1.1. We are done by 2.3.5. □

Jan. 24, 2024

2.5.11 Theorem. *If the set F is a family of sets, then $\bigcup F$ is a set.*

Proof. Let F be built at stage Σ (Princ. 1). Now,

$$x \in \bigcup F \equiv \underset{\text{before } \Sigma}{x} \in \underset{\text{before } \Sigma}{\overset{\text{some}}{\downarrow}\ A} \in \overset{\text{before } \Sigma}{\downarrow}\ F$$

Thus x is available or built *before* stage Σ at which F was built.

x being arbitrary, all members of $\bigcup F$ are available/built *before* Σ, so we can build $\bigcup F$ as a set at stage Σ. □

2.5.12 Remark. What if $\mathbb{F} = \emptyset$? Does it affect Theorem 2.5.10? Yes, **badly**!

In Definition 2.5.7 we read

$$\bigcap \mathbb{F} \overset{Def}{=} \left\{ x : \text{for all } A, \overbrace{\underbrace{A \in \mathbb{F}}_{\mathbf{t}} \to x \in A}^{\mathbf{f}} \right\} \qquad (**)$$

However, as *the hypothesis (i.e., lhs) of the implication in* $(**)$ *is* ***false***, the implication itself is **true**. Thus the entrance condition "for all $A, A \in \mathbb{F} \to x \in A$" is true for *all* x and thus allows ALL objects x to get into $\bigcap \mathbb{F}$,

This means $\bigcap \mathbb{F} = \mathbb{U}$, the universe of *all* objects which we saw (cf. Section 2.2) is a proper class —i.e., *not* a set. □

2.5.13 Exercise. What is $\bigcup F$ if $F = \emptyset$? Set or proper class? Can you "compute" which class it is exactly? □

2.5.14 Remark. (More notation)

Suppose the family of sets Q is a *set* of sets A_i, for $i = 1, 2, \ldots, n$ where $n \geq 3$.

$$Q = \{A_1, A_2, \ldots, A_n\}$$

Then we have a few alternative *notations* for $\bigcap Q$:

(a)

$$A_1 \cap A_2 \cap \ldots \cap A_n$$

or, more elegantly,

(b)

$$\bigcap_{i=1}^{n} A_i$$

or also

(c)

$$\bigcap_{i=1}^{n} A_i$$

Similarly for $\bigcup Q$:

(i)

$$A_1 \cup A_2 \cup \ldots \cup A_n$$

or, more elegantly,

(ii)

$$\bigcup_{i=1}^{n} A_i$$

or also

(iii)

$$\bigcup_{i=1}^{n} A_i$$

If the family has so many elements that *all the natural numbers are needed* to index the sets in the set family Q we will write

$$\bigcap_{i=0}^{\infty} A_i$$

or

$$\bigcap_{i=0}^{\infty} A_i$$

or

$$\bigcap_{i\geq 0} A_i$$

or

$$\bigcap_{i\geq 0} A_i$$

for $\bigcap Q$ and

$$\bigcup_{i=0}^{\infty} A_i$$

or

$$\bigcup_{i=0}^{\infty} A_i$$

or

$$\bigcup_{i\geq 0} A_i$$

or

$$\bigcup_{i\geq 0} A_i$$

for $\bigcup Q$ □

2.5.15 Example. Thus, for example, $A \cup B \cup C \cup D$ can be seen — just changing the notation— as $A_1 \cup A_2 \cup A_3 \cup A_4$, therefore it means, $\bigcup\{A_1, A_2, A_3, A_4\}$, or $\bigcup\{A, B, C, D\}$.

Same comment for $\cap$. □

Pause. How come for the case for $n = 2$ we *proved*[†] $A \cup B = \bigcup\{A, B\}$ (2.5.9) but *here* we say ($n \geq 3$) that something like the content of the previous remark and example are just *notation* (*definitions*)?

Well, we had *independent* definitions (and associated theorems re set status for each, 2.4.2 and 2.5.11) for $A \cup B$ and $\bigcup\{A, B\}$ so it makes sense to compare the two *independent* definitions after the fact and see if we can *prove* that *they say the same thing*.

> For $n \geq 3$ we opted to *NOT* give a definition for $A_1 \cup \ldots \cup A_n$ that is *independent* of $\bigcup\{A_1 \cup \ldots \cup A_n\}$, rather we gave the definition of the former in terms of the latter.

No independent definitions, no theorem to compare the two!◀

[†]Well, *you* proved! Same thing :-)

Chapter 3

The Ordered Pair and Cartesian Products

To introduce the concepts of cartesian product —so that, in principle, plane analytic geometry can be developed within set theory— we need an object "(A, B)" that is *like* the set pair (2.3.1) in that it contains *two* objects, A and B ($A = B$ is a possibility), but in (A, B) order *and* length (here it is 2) matter!

That is,

> *We want* $(A, B) = (A', B')$ implies $A = A'$ *and* $B = B'$. *Moreover,* (A, A) *is not* $\{A\}$*! It is still an ordered* pair *(length = 2) but so happens that the first and second* component *—as we call the members of the ordered* pair*— are equal in this example.*

So, are we going to accept a new type of object in set theory? *Not at all*!

We will *build* (A, B) so that it is a set!

3.0.1 Definition. (Ordered pair) *By definition (Kuratowski)*, (A, B) is the *abbreviation* (short name) given below:

$$(A, B) \stackrel{Def}{=} \Big\{A, \{A, B\}\Big\} \tag{1}$$

We call "(A, B)" an *ordered pair*, and A its first *component*, while B is its second component. □

3.0.2 Remark.

1. Note that $A \neq \{A, B\}$ because we would otherwise get

   $$\text{the right } A \in \text{the left } A$$

   which is false for *sets or atoms* A. Thus (A, B) does contain exactly two *members*, or *has length 2*:

   $$A \text{ and } \{A, B\}.$$

   Pause. We have *not* said in 3.0.1 that A and B are sets or atoms. So what right do we have in the paragraph above to so declare?◀

2. What about the desired property that

   $$(A, B) = (X, Y) \to A = X \wedge B = Y \tag{2}$$

   Well, **assume the lhs** of "$\to$" in (2) and prove the rhs, "$A = X \wedge B = Y$".

   From our truth table we know that we do the latter by proving *each* of $A = X$ and $B = Y$ true (*separately*).

The lhs that we *assumed* translates to

$$\Big\{A, \{A, B\}\Big\} = \Big\{X, \{X, Y\}\Big\} \tag{3}$$

By the remark #1 above there are *two* distinct members in each of the two sets that we equate in (3).

So since (3) is true (by assumption) we have (by definition of set equality) one of:

(a) $A = \{X, Y\}$ and $\{A, B\} = X$, that is, **1st listed element in lhs of "=" equals the 2nd listed in rhs; and 2nd listed element in lhs of "=" equals the 1st listed in rhs**.

(b) $A = X$ and $\{A, B\} = \{X, Y\}$.

Now case (a) above *cannot hold*, for it leads to $A = \{\overbrace{\{A, B\}}^{replaces\ X}, Y\}$. This in turn leads to

$$\{A, B\} \in A$$

and thus the set $\{A, B\}$ is built *before* ONE of its members, A, which contradicts Principle 0.

Jan. 26, 2024

Let's then work with case (b).

We have

$$\{A, B\} = \{A, Y\} \tag{4}$$

Well, all the members on the lhs must also be on the rhs. I note that A is. I have two subcases.

- What if B is also equal to A? Then (4) becomes $\{B\} = \{A, Y\}$ and thus $Y \in \{B\}$ (why?). Hence $Y = B$.

  We showed so far $A = X$ (listed in case (b)) *and* $B = Y$ (proved in this subcase); great!

- In the 2nd and final subcase (Why "final"?) B is *not* equal to A.

  But B must be in the rhs of (4), so the only way —since $A \neq B$— is $B = Y$. *All Done*! □

Worth *recording* as a theorem what we proved above:

3.0.3 Theorem. *If* $(A, B) = (X, Y)$, *then* $A = X$ *and* $B = Y$.

But is (A, B) a set? (atom it is not, of course!) Yes!

3.0.4 Theorem. (A, B) *is a set.*

Proof. Now $(A, B) = \Big\{A, \{A, B\}\Big\}$. By 2.3.1, $\{A, B\}$ is set. Applying 2.3.1 once more, $\Big\{A, \{A, B\}\Big\}$ is a set. □

3.0.5 Example. So, $(1,2) = \{1,\{1,2\}\}$, $(1,1) = \{1,\{1\}\}$, and $(\{a\},\{b\}) = \{\{a\},\{\{a\},\{b\}\}\}$. □

3.0.6 Remark. We can extend the ordered pair to ordered *triple*, ordered *quadruple*, and beyond!

We take this approach in these notes:

$$(A,B,C) \overset{Def}{=} \Big((A,B),C\Big) \tag{1}$$

$$(A,B,C,D) \overset{Def}{=} \Big((A,B,C),D\Big) \tag{2}$$

$$(A,B,C,D,E) \overset{Def}{=} \Big((A,B,C,D),E\Big) \tag{3}$$

ETC. So suppose we defined what an n-tuple is, for *some fixed unspecified* n, and denote it by $(A_1, A_2, \ldots, A_n)$ for convenience.

Then we define $(n+1)$-tuple, in general, by

$$(A_1, A_2, \ldots, A_n, A_{n+1}) \overset{Def}{=} \Big((A_1, A_2, \ldots, A_n), A_{n+1}\Big) \tag{$*$}$$

This is an "*inductive*" or "*recursive*" definition, defining a concept ($n+1$-tuple) in terms of *a smaller instance of itself*, namely, in terms of the concept for an n-tuple, and in terms of the case $n = 2$ that we dealt with by *direct* definition (*not* in terms of the concept itself!) in 3.0.1.

> $(*)$ is a general (for each length n that is) formation rule that allows us to build a tuple *longer by ONE*, as is compared to a tuple *we have already built*.

Suffice it to say this "case of $n+1$ in terms of case of n" provides just *shorthand notation* to take the mystery out of the red capitalised "etc." above. We **condense**/*codify* infinitely many definitions (1), (2), (3), ... into just **two**:

- 3.0.1

and

- $(*)$

The reader has probably seen such recursive definitions before (likely in calculus and/or high school).

The most frequent example that occurs is to define, for any natural number n and any real number $a > 0$, what a^n means. One goes like this:

$a^0 = 1$
$a^{n+1} = a \cdot a^n$

The above condenses *infinitely many definitions* such as

$a^0 = 1$
$a^1 = a \cdot a^0 = a$
$a^2 = a \cdot a^1 = a \cdot a$
$a^3 = a \cdot a^2 = a \cdot a \cdot a$
$a^4 = a \cdot a^3 = a \cdot a \cdot a \cdot a$
$\vdots$

into just two!

We will study *inductive definitions* and *induction* later in the course!

Before we exit this remark note that $(A, B, C) = (A', B', C')$ implies $A = A', B = B', C = C'$ **because** the hypothesis says (3.0.6 (1))

$$((A, B), C) = ((A', B'), C')$$

and thus (3.0.3) implies

$$C = C' \text{ and } (A, B) = (A', B')$$

The second equality implies (3.0.3 again) $A = A'$ and $B = B'$.

That is, (A, B, C) **is** an **ordered** triple (3-tuple).

We can also prove that $(A_1, A_2, \ldots, A_n, A_{n+1})$ is an **ordered** $n+1$-tuple, i.e.,

$$(A_1, A_2, \ldots, A_{n+1}) = (A'_1, A'_2, \ldots, A'_{n+1}) \rightarrow A_1 = A'_1 \wedge \ldots \wedge A_{n+1} = A'_{n+1}$$

IF we have followed the "etc." all the way to the case of $(A_1, A_2, \ldots, A_n)$.

We will do the "etc."-argument *elegantly* once we learn induction!

□ 

3.0.7 Definition. (Finite sequences) An n-tuple for $n \geq 1$ is called a finite sequence of length n, where we extend the concept to a *one element sequence* —**by definition**— to be

$$(A) \overset{Def}{=} A$$

□

Note that now we can redefine all sequences of lengths $n \geq 1$ —pushing the starting point of the "etc.-construction" in 3.0.6 to $n = 1$ (from $n = 2$).

Using again $(*)$ above, but this time with starting condition that of 3.0.7, for $n = 2$ we rediscover (A_1, A_2):

$$\text{the "new" 2-tuple pair: } (A_1, A_2) \overset{\text{by } (*)}{=} \Big((A_1), A_2\Big) \overset{\text{by } 3.0.7}{=} \Big(A_1, A_2\Big)$$

The big red brackets are applications of the ordered pair defined in 3.0.1, just as it was in the general definition $(*)$.

3.1. The Cartesian product

Jan. 29, 2024

We next define classes of *ordered* pairs.

3.1.1 Definition. (Cartesian product of classes) Let $\mathbb{A}$ and $\mathbb{B}$ be classes. Then we define

$$\mathbb{A} \times \mathbb{B} \overset{Def}{=} \Big\{(x, y) : x \in \mathbb{A} \wedge y \in \mathbb{B}\Big\}$$

The definition requires both sides of $\times$ to be classes. It makes no sense if one or both are atoms.

□

3.1.2 Theorem. *If A and B are sets, then so is $A \times B$.*

Proof. By 3.1.1 and 3.0.1

$$A \times B = \Big\{ \{x, \{x, y\}\} : x \in A \wedge y \in B \Big\} \tag{1}$$

Plan: I want to "find" a *set* "X" so that the inclusion $A \times B \subseteq X$ is true. Then I can apply the *subclass theorem* (2.3.5).

Thus I am starting my search with "let $\{x, \{x, y\}\} \in A \times B$" and I am analysing this statement attempting to find an X such that $\{x, \{x, y\}\} \in X$.

So, for each $\{x, \{x, y\}\} \in A \times B$ we have $x \in A$ and $\{x, \overset{\text{in } B}{\overset{\downarrow}{y}}\} \subseteq A \cup B$, or $x \in A$ and $\{x, y\} \in 2^{A \cup B}$.

Thus $\{x, \{x, y\}\} \subseteq A \cup 2^{A \cup B}$ and hence (changing notation)

$$(x, y) \in 2^{A \cup 2^{A \cup B}} \tag{2}$$

I found *a SET* —"$X = 2^{A \cup 2^{A \cup B}}$"— that works, *meaning*

$$A \times B \subseteq X$$

We have established —by the arbitrariness of x, y and by (2)— that

$$A \times B \subseteq 2^{A \cup 2^{A \cup B}}$$

thus $A \times B$ is a set by 2.3.5, 2.4.2 and 2.5.2. □

3.1.3 Definition. Mindful of the Remark 3.0.6 where we defined $\Big(A, B, C\Big)$ as short for $\Big((A, B), C\Big)$, $\Big(A, B, C, D\Big)$ as short for $\Big((A, B, C), D\Big)$, etc., we define here $A_1 \times \ldots \times A_n$ for any $n \geq 3$ to mean

$$\{(x_1, x_2, \ldots x_n) : x_i \in A_i, \text{ for } i = 1, \ldots, n\}$$

and then remark as follows: Thus,

$$\begin{aligned}
A \times B \times C &\overset{\text{Def of } A\times B\times C}{=} \{(x, y, z) : x \in A \wedge y \in B \wedge z \in C\} \\
&= \{((x, y), z) : x \in A \wedge y \in B \wedge z \in C\} \\
&= \{((x, y), z) : (x, y) \in A \times B \wedge z \in C\} \\
&= (A \times B) \times C \\
A \times B \times C \times D &\overset{Def}{=} (A \times B \times C) \times D \\
&\vdots \\
A_1 \times A_2 \times \ldots \times A_n \times A_{n+1} &\overset{Def}{=} \Big\{(x_1, x_2, \ldots, x_n, x_{n+1}) : x_i \in A_i\Big\} \\
&= \{((x_1, \ldots, x_n), x_{n+1}) : x_i \in A_i\} \\
&= \{((x_1, \ldots, x_n), x_{n+1}) : (x_1, \ldots, x_n) \in \\
&\qquad (A_1 \times A_2 \times \cdots \times A_n) \wedge x_{n+1} \in A_{n+1}\} \\
&= (A_1 \times \cdots A_n) \times A_{n+1} \\
&\vdots
\end{aligned}$$

$$\text{We may write } \bigtimes_{i=1}^{n} A_i \text{ for } A_1 \times A_2 \times \ldots \times A_n$$

If $A_1 = \ldots = A_n = B$ we may write B^n for $A_1 \times A_2 \times \ldots \times A_n$. □

3.1.4 Remark. Thus, what we learnt in 3.1.3 is, in other words,

$$\bigtimes_{i=1}^{n} A_i \overset{Def}{=} \Big\{(x_1, \ldots, x_n) : x_i \in A_i, \text{ for } i = 1, 2, \ldots, n\Big\}$$

and

$$B^n \overset{Def}{=} \Big\{(x_1, \ldots, x_n) : x_i \in B\Big\}$$

□

3.1.5 Theorem. *If A_i, for $i = 1, 2, \ldots, n$ is a set, then so is $\bigtimes_{i=1}^{n} A_i$.*

Proof. $A \times B$ is a set by 3.1.2. By 3.1.3, **and in this order**, we verify that so is $A \times B \times C$* and $A \times B \times C \times D$ and ... and $A_1 \times A_2 \times \ldots \times A_n$ and ... □

If we had inductive definitions available already, then Definition 3.1.3 would simply read

$$A_1 \times A_2 \overset{Def}{=} \Big\{(x_1, x_2) : x_1 \in A_1 \wedge x_2 \in A_2\Big\}$$

and, for $n \geq 2$,

$$A_1 \times A_2 \times \ldots \times A_n \times A_{n+1} \overset{Def}{=} (A_1 \times A_2 \times \ldots \times A_n) \times A_{n+1}$$

Correspondingly, the proof of 3.1.5 would be far more elegant, via induction.

*Because $A \times B \times C = (A \times B) \times C$.

Chapter 4

Relations and functions

The topic of relations and functions is central in all *mathematics* and *computing*.

In *mathematics*, whether it is calculus, algebra or anything else, one deals with relations (notably *equivalence relations*, *order*) and all sorts of functions while in *computing* one computes relations and functions, that is, writing programs that given an input to a relation they compute the response (true or false) or given an input to a function they compute a response which is some object (number, graph, tree, matrix, other) or *nothing*, *in case there is no response* for said input (for example, there is no response to input "x, y" if what we are computing is $\frac{x}{y}$ but $y = 0$).

We are taking an "extensional" point of view in this course —as is customary in set theory, algebra, calculus and discrete math— of relations and functions, that is, *we view them as classes of (input, output) ordered* **pairs**.

It is also possible to take an *intentional* point of view, *especially in computer science* and some specific areas of mathematics, viewing relations and functions as *methods* to compute outputs from given inputs.

4.1. Relations

4.1.1 Definition. (Binary relation) A binary relation is a class $\mathbb{R}$[†] of ordered pairs.

> The statements $(x, y) \in \mathbb{R}$, $x\mathbb{R}y$ and $\mathbb{R}(x, y)$ are *equivalent; that is,* **they mean the same thing.**

$x\mathbb{R}y$ is the preferred "*infix*" notation —imitating notation such as $A \subset B$, $x < y$, $x = y$ and has notational advantages. □

4.1.2 Remark. $\mathbb{R}$ contains just pairs (x, y), that is, just *sets* $\{x, \{x, y\}\}$, that is, it is a *family of sets*.

> Since $(x_1, x_2, \ldots, x_n) = \Big((x_1, x_2, \ldots, x_{n-1}), x_n\Big)$, it follows that binary relations (classes of ordered pairs) *is ALL we need to study*.

[†]I write "$\mathbb{R}$" or "R" for a relation, generically, but $\mathbb{P}$, $\mathbb{Q}$, $\mathbb{S}$ are available to use as well. I will avoid specific names such as $<$, $\subseteq$ in a general discussion. These two are apt to bring in in examples.

BTW, a class of ordered n-tuples, $(x_1, x_2, \ldots, x_n)$, is called *an n-ary relation*. As I said above we do not need to pay special attention to them. □ 

Jan. 31, 2024

4.1.3 Example. Examples of relations:

(i) $\emptyset$ Since this set contains nothing I can imagine that it is a set of zero number of pairs.

(ii) $\{(1,1)\}$

(iii) $\{(1,1),(1,2)\}$

(iv) $\mathbb{N}^2$, that is $\{(x,y) : x \in \mathbb{N} \wedge y \in \mathbb{N}\}$. This is a set by the fact that $\mathbb{N}$ is (Why?) and thus so is $\mathbb{N} \times \mathbb{N}$ by 3.1.2.

(v) $<$ on $\mathbb{N}$, that is $\{(x,y) : x < y \wedge x \in \mathbb{N} \wedge y \in \mathbb{N}\}$. This is a set since $< \subseteq \mathbb{N}^2$.

(vi) $\in$, that is,

$$\{(x,y) : x \in y \wedge x \in \mathbb{U} \wedge y \in \mathbb{V}\} \qquad (*)$$

This is a *proper* class (non set). Why? Well,

(a) If $\in$ is a *set* then so is *its SUBclass*

$$\{(x,\{x\}) : x \in \mathbb{U}\} = \left\{\left\{x,\{x,\{x\}\}\right\} : x \in \mathbb{U}\right\} \qquad (**)$$

(b) By the Union Theorem 2.5.11

$$\bigcup\left\{\left\{x,\{x,\{x\}\}\right\} : x \in \mathbb{U}\right\} = \left\{x, x'', x''', \ldots, \{x,\{x\}\}, \{x',\{x'\}\}, \{x'',\{x''\}\}\ldots\right\}$$

is a *set*. This "set" has $\mathbb{U}$ as a subclass (due to the "loose" $x, x', x'', \ldots$) contradicting the subclass theorem.

□

So, a binary relation $\mathbb{R}$ is a table of pairs:

Table 4.1:

input: x	output: y
a	b
a'	b'
$\vdots$	$\vdots$
u	v
$\vdots$	$\vdots$

1. Thus one way to view R is as a device that for inputs x, valued $a, a', \ldots, u, \ldots$ one gets the outputs y, valued $b, b', \ldots, v, \ldots$ respectively. It is all right that a given input may yield *multiple* outputs (e.g., case (iii) in the previous example).

2. Another point of view is to see *both* x and y as inputs of R and the outputs then are **true** or **false** (**t** or **f**). *Such is the way we often view the relations $<$ and $=$ on the natural numbers.*

   For example, (a, b) is in the table (that is, aRb is true) hence if both a and b are ordered input values, then the relation outputs **t**.

Most of the time we will take the point of view in 1 above. This point of view compels us to define *domain* and *range* of a relation $\mathbb{R}$, that is, the class of all inputs that *cause an output* and the class of all *caused outputs* respectively.

4.1.4 Definition. (Domain and range) For any relation $\mathbb{R}$ we define *domain*, in symbols "dom" by

$$\text{dom}(\mathbb{R}) \stackrel{Def}{=} \{x : (\exists y) x \mathbb{R} y\}$$

where we have introduced the notation "$(\exists y)$" as short for "there exists some y such that", or "for some y".

Range, in symbols "ran", is defined also in the obvious way:

$$\text{ran}(\mathbb{R}) \stackrel{Def}{=} \{x : (\exists y) y \mathbb{R} x\}$$ □

Thus the domain of $\mathbb{R}$ is the class containing *precisely all the entries* of the **left column** of Table 4.1 on p.87 while the range contains *precisely all the entries* of the **right column**.

We settle the following, before other things:

4.1.5 Theorem. *For a* set *relation* R*, both* $\mathrm{dom}(R)$ *and* $\mathrm{ran}(R)$ *are sets.*

Proof. For **domain** we collect *ALL* the x such that xRy, for some y, that is, all the x such that

$$\{x, \{x, y\}\} \in R \tag{1}$$

for some y.

So, R is a *set* family of sets

$$\Big\{\{x, \{x, y\}\}, \{x', \{x', y'\}\}, \{x'', \{x'', y''\}\}, \dots\Big\}$$

Thus, taking the family union, I have

$$\Big\{x, \{x, y\}, x', \{x', y'\}, x'', \{x'', y''\}, \dots\Big\} = \bigcup R$$

and $\mathrm{dom}(R)$ is the collection of all the "loose" $x, x', x'', \dots$ above (4.1.4).

Therefore

$$\mathrm{dom}(R) \subseteq \bigcup R \tag{\dagger}$$

Now, R is a set-family of sets, thus $\bigcup R$ is a set. But then by (†) and the **subclass theorem**, $\mathrm{dom}(R)$ is a set. This settles the domain case.

Let $\mathcal{A}$ be the set of *ALL atoms* (anywhere).

Pause. Why is the class of *all atoms* a *set*?◀

Now define

$$S \overset{Def}{=} \Big(\bigcup R\Big) - \mathcal{A}$$

So, S is a *set family* —we just **removed** *all atom members* of $\bigcup R$— and it contains *all* the $\{x, y\}$ parts of *all* $\{x, \{x, y\}\} \in R$. Thus,

$$S = \Big\{\{x, y\}, \{x', y'\}, \{x'', y''\}, \dots;\ \text{plus those } x, x', x'', \dots \text{ that are } \textit{sets}\Big\}$$

Then $\bigcup S$ contains all the y (and other things). That is, $\text{ran}(R) \subseteq \bigcup S$, and this settles the range case. □

4.1.6 Exercise. Armed with the theorem 4.1.5 above revisit the **relation** $\in$ and easily prove that it is a proper class (not a set relation). □

4.1.1. Fields

Feb. 2, 2024

4.1.7 Definition. In practice we often have an *a priori decision* about what are *in principle* **"legal" inputs** for a relation $\mathbb{R}$, and where its outputs go.

Thus we have two classes, $\mathbb{A}$ and $\mathbb{B}$ for the class of legal inputs and possible outputs respectively. Clearly we have $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{B}$.

We call $\mathbb{A}$ and $\mathbb{B}$ left field and right field respectively, and instead of $\mathbb{R} \subseteq \mathbb{A} \times \mathbb{B}$ we often write

$$\mathbb{R} : \mathbb{A} \to \mathbb{B}$$

and also

$$\mathbb{A} \xrightarrow{\mathbb{R}} \mathbb{B}$$

pronounced "$\mathbb{R}$ is a relation *from* $\mathbb{A}$ *to* $\mathbb{B}$".

Thus, "Let $\mathbb{A} \xrightarrow{\mathbb{R}} \mathbb{B}$", in proper English, says "**Let $\mathbb{R}$ be a relation with left field $\mathbb{A}$ and right field $\mathbb{B}$**".

The term *field* —without left/right qualifiers— for $\mathbb{R} : \mathbb{A} \to \mathbb{B}$ refers to $\mathbb{A} \cup \mathbb{B}$.

If $\mathbb{A} = \mathbb{B}$ then we have

$$\mathbb{R} : \mathbb{A} \to \mathbb{A}$$

but rather than pronouncing this as "$\mathbb{R}$ is a relation *from* $\mathbb{A}$ *to* $\mathbb{A}$" we *prefer*[†] to say "$\mathbb{R}$ is on $\mathbb{A}$". □

[†]Both ways of saying it are correct.

4.1.8 Example. The *a priori* legal inputs in *Number Theory* and in *Computability* are all the natural numbers from $\mathbb{N}$.

In calculus inputs are real (from $\mathbb{R}$) and so are outputs (in $\mathbb{R}$). But it is not the case that all inputs cause outputs! There is no (real) output for $\sqrt{-2.5}$ and there is no Natural Number output from $\sqrt{n}$ for all n —take for example $n = 2$. □

You will pardon —I hope— the use of $\mathbb{R}$ for a generic relation but also for *the set of all reals*.

4.1.9 Remark. Trivially, for any $\mathbb{R} : \mathbb{A} \to \mathbb{B}$, we have $\text{dom}(\mathbb{R}) \subseteq \mathbb{A}$ and $\text{ran}(\mathbb{R}) \subseteq \mathbb{B}$. To see this think of 4.1 and its columns representing $\text{dom}(\mathbb{R})$ and $\text{ran}(\mathbb{R})$.

4.1.10 Exercise. Give a quick proof of each of the above inclusions. □

Also, for any relation $\mathbb{P}$ with no **a priori** specified left/right fields, $\mathbb{P}$ is a relation from $\text{dom}(\mathbb{P}) \to \text{ran}(\mathbb{P})$.

Naturally, we say that $\text{dom}(\mathbb{P}) \cup \text{ran}(\mathbb{P})$ is the *field* of $\mathbb{P}$ in this case. □

4.1.2. Totalness and Ontoness

4.1.11 Example. As an example, consider the *divisibility relation* on all integers (their set denoted by $\mathbb{Z}$) denoted by "|":

$$x|y \text{ means } x \text{ divides } y \text{ with 0 remainder}$$

thus, for $x = 0$ and all y, the division is *illegal*, therefore

> *The input $x = 0$ to the relation "|"* ***produces no output****, in other words,* ***"for input $x = 0$ the relation is undefined."***

We walk away with two things from this example:

1. It **does** make sense for some relations to *a priori* choose left and right fields, here
$$| : \mathbb{Z} \to \mathbb{Z}$$
You would not have divisibility on *real numbers*!

2. dom(|) is the set of all inputs that **produce** some output. Thus, it is NOT the case for all relations that their domain is the same as the left field *chosen*! Note the case in this example! And forget the term "codomain" that you may find in fake publications on discrete MATH out there! □ 

4.1.12 Example. Next consider the relation $<$ with left/right fields restricted to $\mathbb{N}$. Then $\text{dom}(<) = \mathbb{N}$, but $\text{ran}(<) \subsetneqq \mathbb{N}$. Indeed, $0 \in \mathbb{N} - \text{ran}(<)$. □

Let us extract some terminology from the above examples:

4.1.13 Definition. Given

$$\mathbb{R} : \mathbb{A} \to \mathbb{B}$$

If $\text{dom}(\mathbb{R}) = \mathbb{A}$, then we call $\mathbb{R}$ *total* or *totally defined*. If $\text{dom}(\mathbb{R}) \subsetneqq \mathbb{A}$, then we say that $\mathbb{R}$ is *nontotal*.

If $\text{ran}(\mathbb{R}) = \mathbb{B}$, then we call $\mathbb{R}$ *onto*. If $\text{ran}(\mathbb{R}) \subsetneqq \mathbb{B}$, then we say that $\mathbb{R}$ is *not onto*. □

So, the relation $|$ above is *nontotal*, and $<$ is *not* onto.

4.1.14 Example. Let $A = \{1, 2\}$.

- The relation $\{(1, 1)\}$ on A is neither total nor onto.
- The relation $\{(1, 1), (1, 2)\}$ on A is onto but not total.
- The relation $\{(1, 1), (2, 1)\}$ on A is total but not onto.
- The relation $\{(1, 1), (2, 2)\}$ on A is total *and* onto.
- The relation $\{(1, 2), (2, 1)\}$ on A is total *and* onto. □

4.1.3. Diagonal or Identity and other Special Types of Relations

4.1.15 Definition. The relation Δ_A on the set A is given by

$$\Delta_A \stackrel{Def}{=} \{(x, x) : x \in A\}$$

We call it the *diagonal* ("Δ" for "diagonal") or *identity* relation on A.

Consistent with the second terminology, we may also use the symbol $\mathbf{1}_A$ for this relation. □

4.1.16 Definition. A relation R (not *a priori* restricted to have *predetermined* left or right fields) is

1. *Transitive*: Iff $xRy \wedge yRz$ implies xRz.
2. *Symmetric*: Iff xRy implies yRx.
3. *Antisymmetric*: Iff $xRy \wedge yRx$ implies $x = y$.
4. *Irreflexive*: Iff xRy implies $x \neq y$. Also said this way: *For NO x can we have xRx*.
5. Now assume R is *on a set* A. Then we call it reflexive iff $\Delta_A \subseteq R$.

□

4.1.17 Example.

(i) *Transitive* examples: $\emptyset$ (*vacuously*), $\{(1,1)\}$, $\{(1,2),(2,3),(1,3)\}$, $<$, $\leq$, $=$, $\mathbb{N}^2$.

(ii) *Symmetric* examples: $\emptyset$ (*vacuously*), $\{(1,1)\}$, $\{(1,2),(2,1)\}$, $=$, $\mathbb{N}^2$.

(iii) *Antisymmetric* examples: $\emptyset$ (*vacuously*), $\{(1,1)\}$, $=$, $\leq$, $\subseteq$.

(iv) *Irreflexive* examples: $\emptyset$ (*vacuously*), $\{(1,2)\}$, $\subsetneq$, the relations "$<$" and "$\neq$" on $\mathbb{N}$.

(v) *Reflexive* examples: $\mathbf{1}_A$ on A, $\{(1,1)\}$ on $\{1\}$, $\{(1,2),(2,1),(1,1),(2,2)\}$ on $\{1,2\}$, $=$ on $\mathbb{N}$, $\leq$ on $\mathbb{N}$. □

4.2. Relational Composition

We can compose relations:

4.2.1 Definition. (Relational composition) Let $\mathbb{R}$ and $\mathbb{S}$ be (possibly NON set) relations.

Then, their **composition**, *in that order*, denoted by $\mathbb{R} \circ \mathbb{S}$ is defined for all x and y by:

$$x\mathbb{R} \circ \mathbb{S}y \overset{Def}{\equiv} (\exists z)\Big(x\mathbb{R}z \wedge z\mathbb{S}y\Big)$$

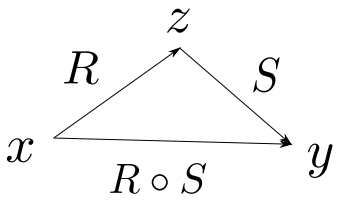
It is *customary* (lazy and incorrect, though) to *abuse* notation and write "$x\mathbb{R}z\mathbb{S}y$" for "$x\mathbb{R}z \wedge z\mathbb{S}y$" just as one writes $x < y < z$ for $x < y \wedge y < z$. □

4.2.2 Example. (Important) Here is whence the emphasis "*in that order*" above. Say, $R = \{(1,2)\}$ and $S = \{(2,1)\}$. Thus, $R \circ S = \{(1,1)\}$ while $S \circ R = \{(2,2)\}$. Thus, $R \circ S \neq S \circ R$ *in general*. □

4.2.3 Example. For any R, we *diagrammatically* indicate xRy by

$$x \xrightarrow{R} y$$

Thus, the situation where we have that $xR \circ Sy$ means, for some z, $xRzSy$ is depicted as:



□ 

4.2.4 Theorem. (Associativity of composition) *For any relations* $\mathbb{R}, \mathbb{S}$ *and* $\mathbb{T}$*, we have*

$$(\mathbb{R} \circ \mathbb{S}) \circ \mathbb{T} = \mathbb{R} \circ (\mathbb{S} \circ \mathbb{T})$$

We state and prove this central result for any class relations.

Proof. We have two directions:

$\rightarrow$: Fix x and y and *let* $x(\mathbb{R} \circ \mathbb{S}) \circ \mathbb{T}y$.

Then, for some z, we have $x(\mathbb{R} \circ \mathbb{S})z\mathbb{T}y$ and hence for some w, the above becomes

$$x\mathbb{R}w\mathbb{S}z\mathbb{T}y \tag{1}$$

But $w\mathbb{S}z\mathbb{T}y$ means $w\mathbb{S} \circ \mathbb{T}y$

hence we rewrite (1) as

$$x\mathbb{R}w(\mathbb{S} \circ \mathbb{T})y$$

Finally, the above says $x\mathbb{R} \circ (\mathbb{S} \circ \mathbb{T})y$.

$\leftarrow$: Just as the $\rightarrow$ case; read if you wish.

Fix x and y and let $x\mathbb{R} \circ (\mathbb{S} \circ \mathbb{T})y$.

Then, for some z, we have $x\mathbb{R}z(\mathbb{S} \circ \mathbb{T})y$ and hence for some u, the above becomes

$$x\mathbb{R}z\mathbb{S}u\mathbb{T}y \tag{2}$$

But $x\mathbb{R}z\mathbb{S}u$ means $x\mathbb{R} \circ \mathbb{S}u$, hence we rewrite (2) as

$$x(\mathbb{R} \circ \mathbb{S})u\mathbb{T}y$$

Finally, the above says $x(\mathbb{R} \circ \mathbb{S}) \circ \mathbb{T}y$. □

The following is almost unnecessary, but offered for emphasis:

4.2.5 Corollary. *If R, S and T are (set) relations, all on some set A,*[†] *then "$R \circ S \circ T$" has a meaning* independent of how brackets are inserted.

The corollary allows us to just omit brackets in a chain of compositions, even longer than the above. It also leads to the definition of relational exponentiation, below:

4.2.6 Definition. (Powers of a binary relation) Let R be a (set) relation. We define R^n, for $n > 0$, as

$$\underbrace{R \circ R \circ \cdots \circ R}_{n \ R} \tag{1}$$

Note that the resulting relation in (1) is independent of how brackets are inserted (4.2.5). It depends only on R and n.

If moreover we have defined R to be *on a set A*, then we also define the 0-th power: R^0 stands for Δ_A or $\mathbf{1}_A$. □

[†]Recall that "R is on a set A" means $R \subseteq A^2$, which is the same as $R : A \to A$.

Feb. 4, 2024

4.2.7 Theorem. *The composition of two (set) relations R and S in that order is also a set.*

Proof. Trivially, $\boxed{R \circ S \subseteq \text{dom}(R) \times \text{ran}(S)}$

because

1. IF $(x, y) \in R \circ S$, that is, $xR \circ Sy$

2. THEN

$$\overbrace{x}^{x \in \text{ dom}(R)\ 4.1.4} \quad R \quad z \quad S \quad \overbrace{y}^{y \in \text{ ran}(S)\ 4.1.4}$$

$$\boxed{\textit{Hence } (x, y) \in \text{dom}(R) \times \text{ran}(S).}$$

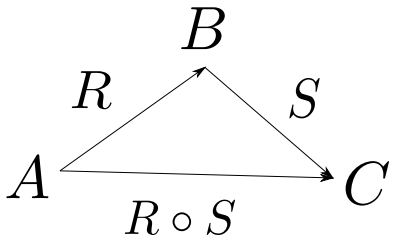
Moreover, we proved in 4.1.5 that $\text{dom}(R)$ and $\text{ran}(S)$ are sets. Thus so is $\text{dom}(R) \times \text{ran}(S)$ (3.1.2).

The boxed formula is proved by 2.3.5. □

4.2.8 Corollary. *If we have $R : A \to B$ and $S : B \to C$, then $R \circ S : A \to C$.*

Proof. From $R \circ S \subseteq \text{dom}(R) \times \text{ran}(S)$ above and $\text{dom}(R) \subseteq A$ and $\text{ran}(S) \subseteq C$. □

The result of the corollary is depicted diagrammatically as



4.2.9 Remark. Say aR^nb.

Then, viewing R^n as $R \circ R^{n-1}$ —$R \circ \overbrace{R \circ \cdots \circ R}^{n-1\ R}$— I have

$$\begin{aligned}
aR^nb &\iff aRa_1R^{n-1}b && \text{for some } a_1\\
&\iff aRa_1Ra_2R^{n-2}b && \text{similarly, for some } a_2\\
&\iff aRa_1Ra_2Ra_3R^{n-3}b && \text{similarly, for some } a_3\\
&\vdots\\
&\iff \overbrace{aRa_1Ra_2Ra_3Ra_4\ldots a_{n-1}\underbrace{R^{n-(n-1)}}_{says\ R}b}^{n\ R} && \text{similarly, for some } a_{n-1}
\end{aligned}$$

Summarising:

Thus aR^nb means that for some $a_1, a_2, \ldots, a_{n-1}$ we have

$$\underset{\cap\\ \mathrm{dom}(R)}{a} \quad Ra_1Ra_2Ra_3Ra_4\ldots a_{n-1}R \quad \underset{\cap\\ \mathrm{ran}(R)}{b} \tag{1}$$

So, $R^n \subseteq \mathrm{dom}(R) \times \mathrm{ran}(R)$.

Alternative proof: I proved above that for a set relation R and $n > 0$ the power R^n is a set. Essentially the above can be condensed as

1. Let xR^ny. This says $xR \circ R^{n-1}y^{\dagger}$ hence $xRzR^{n-1}y$, for some z, thus $x \in \mathrm{dom}(R)$ as implied by xRz.
2. Let again xR^ny. This says $xR^{n-1} \circ Ry^{\ddagger}$ hence $xR^{n-1}wRy$, for some w, thus $y \in \mathrm{ran}(R)$ as implied by wRy.

Compactly, (1) and (2) say $(x, y) \in \mathrm{dom}(R) \times \mathrm{ran}(R)$. So, $R^n \subseteq \mathrm{dom}(R) \times \mathrm{ran}(R)$. □

$^{\dagger}R^n = \overbrace{R \circ R \circ \cdots \circ R}^{n\ times} = R \circ \overbrace{R \circ \cdots \circ R}^{n-1\ times}.$

$^{\ddagger}R^n = \overbrace{R \circ R \circ \cdots \circ R}^{n\ times} = \overbrace{R \circ \cdots \circ R}^{n-1\ times} \circ R.$

4.2.10 Exercise. Let R be a relation on A. Then for all $n \geq 0$, R^n is a set.

Hint. See (1) above.

□

4.2.11 Example. Let $R = \{(1,2),(2,3)\}$. What is R^2?

Well, when can we have xR^2y? Precisely if/when we can find x, y, z that satisfy $xRzRy$. By direct inspection, the values $x = 1$, $y = 3$ and $z = 2$ are the *only ones* that satisfy $xRzRy$.

Thus $1R^23$, or $(1,3) \in R^2$. We conclude $R^2 = \{(1,3)\}$ by the "only ones" above. □

4.2.12 Exercise. Show that if for a relation R we know that $R^2 \subseteq R$, then R is transitive and conversely. □

4.3. Transitive closure

4.3.1 Definition. (Transitive closure of R) Ⓐ *transitive closure* of a relation R —if it exists— is ⓐ $\subseteq$-*smallest* transitive T that *contains R as a subset*.

More precisely,

1. T is transitive, and $R \subseteq T$.
2. If S is also transitive and $R \subseteq S$, then $T \subseteq S$. This makes the term "$\subseteq$-*smallest*" precise. □

Note that we *hedged twice* in the definition, because **at this point** we do not know yet:

- If every relation has a transitive closure; hence the "if it exists".
- We do not know *if it is unique* either, hence the circled indefinite articles "A" and "a".

4.3.2 Remark. *Uniqueness* can be settled immediately *from the definition above*: Suppose T and T' fulfil Definition 4.3.1, that is, suppose *both are* transitive closures of some R. Thus,

1. $R \subseteq T$

   and

2. $R \subseteq T'$

since both are closures.

But now think of T as a closure and T' as the "S" of 4.3.1 (it includes R all right!)

Hence $T \subseteq T'$.

Now reverse the role-playing and think of T' as a closure, while T plays the role of "S". We get $T' \subseteq T$. Hence, $T = T'$. □

4.3.3 Definition. The **unique** transitive closure, *if it exists*, is denoted by R^+. □

4.3.4 Exercise. If R is transitive, then R^+ exists. In fact, $R^+ = R$. □

The above exercise is hardly exciting, but learning that R^+ exists for *every* R and also learning how to "compute" R^+ *is* exciting. We do this next.

Feb. 7, 2024

4.3.5 Lemma. *Given a (set) relation R. Then $\bigcup_{n=1}^{\infty} R^n$ is a transitive (set) relation.*

Proof. We have two things to do.

1. $\bigcup_{n=1}^{\infty} R^n$ is a set.
2. $\bigcup_{n=1}^{\infty} R^n$ is a transitive relation.

Proof of **1.** Since we are using the notation from 2.5.14, we *must* first show that the family

$$\mathbb{F} = \left\{ R, R^2, \dots, R^i, \dots \right\}$$

is a set. **We already know that each R^i, $i \geq 1$, is a set.**

By 4.2.9,

$$R^i \subseteq \text{dom}(R) \times \text{ran}(R)$$

for $i \geq 1$, *OR*

$$R^i \in 2^{\text{dom}(R)\times\text{ran}(R)}$$

for $i \geq 1$.

Therefore

$$\mathbb{F} \subseteq 2^{\text{dom}(R)\times\text{ran}(R)}$$

and hence $\mathbb{F}$ is a *set* family (2.3.5) of sets and we can use the notation from 2.5.14 to write

$$\bigcup_{i=1}^{\infty} R^i = \bigcup \mathbb{F}$$

which is *a set*, as we know (2.5.11).

Proof of **2.** Now, $\bigcup_{i=1}^{\infty} R^i$ is also, of course, a *binary relation* since trivially it is a *set* of *ordered pairs*.

Next, we prove it is *transitive*.

Let

$$x \bigcup_{i=1}^{\infty} R^i \, y \; \bigcup_{i=1}^{\infty} R^i \, z$$

Thus for some n and m we have (*see footnote below*)

$$x \, R^n \, y \, {}^{\dagger}R^m \, z$$

this says the same thing as

$$x \overbrace{R \circ R \circ \cdots R}^{n} \, y \overbrace{R \circ R \circ \cdots R}^{m} \, z$$

or

$$x \overbrace{R \circ R \circ \cdots R}^{n} \circ \overbrace{R \circ R \circ \cdots R}^{m} \, z$$

or

$$x \overbrace{R \circ R \circ \cdots R}^{n+m} \, z$$

Hence, since $(x, z) \in R^{n+m}$ from above, we have

$$(x, z) \in \bigcup \Big\{ \ldots, R^{n+m}, \ldots \Big\}, \text{ that is, (2.5.14), } x \bigcup_{i=1}^{\infty} R^i \, z$$

□

[†] $x \bigcup_{i=1}^{\infty} R^i \, y$ means $(x, y) \in \bigcup_{i=1}^{\infty} R^i$, therefore $(x, y) \in R^n$ for some n by definition of $\bigcup_{n=1}^{\infty}$.

Since $R \subseteq \bigcup_{i=1}^{\infty} R^i$ due to $R = R^1$, all that remains to show that $\bigcup_{i=1}^{\infty} R^i$ is a transitive closure of R is to show the Lemma below.

4.3.6 Lemma. *If $R \subseteq S$ and S is transitive, then $\bigcup_{i=1}^{\infty} R^i \subseteq S$.*

Proof. I will just show *instead* that **for all** $n \geq 1$, $R^n \subseteq S$.

(1) OK, $R \subseteq S$ is our *assumption*, thus $R^1 \subseteq S$ is true.

(2) For $R^2 \subseteq S$ let xR^2y, thus (for some z), $xRzRy$ hence $xSzSy$. But S is transitive, so xSy. Done.

(3) For $R^3 \subseteq S$ let xR^3y, thus (for some z), xR^2zRy hence $\overbrace{xSz}^{By\ (2)}\ Sy$. But S is transitive, so the last expression gives xSy. Done.

$(n+1)$ **You see the pattern**: Pretend we proved up to *some fixed unspecified* n:

$$R^n \subseteq S \qquad (\ddagger)$$

Thus, for the $n+1$ case, for the same n we just fixed,

$$xR^{n+1}y \Leftrightarrow xR^n \circ Ry \Leftrightarrow xR^nzRy \text{ (some } z\text{)} \overset{by\ (\ddagger)}{\Rightarrow} xSzSy \Rightarrow xSy^{\dagger}$$

□

[†] S is transitive.

We have proved:

4.3.7 Theorem. (*The* transitive closure exists) *For any relation R, its transitive closure R^+ exists and is unique. We have that $R^+ = \bigcup_{i=1}^{\infty} R^i$.*

Feb. 9, 2024

An interesting corollary that will lend a computational flavour to 4.3.7 is the following.

4.3.8 Corollary. *If R is on the set $A = \{a_1, a_2, \ldots, a_n\}$ where, for $i = 1, \ldots, n$, the a_i are distinct, then $R^+ = \bigcup_{i=1}^{n} R^i$.*

Proof. By 4.3.7, all we have to do is prove

$$R^+ = \bigcup_{i=1}^{\infty} R^i \subseteq \bigcup_{i=1}^{n} R^i \tag{1}$$

since the $\supseteq$ part is obvious.

So let $x \bigcup_{i=1}^{\infty} R^i y$. This means that

$$xR^q y, \text{ for some } q \geq 1 \tag{2}$$

Thus, I have two cases for (2):

Case 1. $q \leq n$. Then $x \bigcup_{i=1}^{n} R^i y$ since $R^q \subseteq \bigcup_{i=1}^{n} R^i$, R^q being one of the "R^i", $1 \leq i \leq n$.

Case 2. $q > n$. In this case I will show that there is *also* a $k \leq n$ such that $xR^k y$, which sends me back to the "easy **Case 1**".

So, I have

$$xR^q y, \text{ for some } q > n \tag{3}$$

Well, if there is one $q > n$ that satisfies (3) there are probably more. Let us pretend that our q is *the smallest* $\boxed{> n}$ that gives us (3).

Wait! Why IS there a *smallest* q such that (3) holds?

Because among those "q" that fit (3)[†] imagine we fix attention to one such.

Now, if it is not the smallest such out of luck(!), then go down to the *next smaller* one that still satisfies (3), call it "q'" instead of "q".

Now go down to the next smaller, $q'' > n$, if q' is not smallest.

Continue like this. Can I do this forever? That is, can we have the following?

$$n < \ldots < q^{(k)\ddagger} < \ldots < q''' < \ldots < q'' < q' < q$$

If yes, then I will have an infinite "descending" chain of distinct natural numbers between q and n.

Absurd!

So I will "hit" the smallest q that satisfies (3).

[†]There is at least one, else we would **not** be in **Case 2**.
[‡]By "$q^{(k)}$" I mean q with k primes.

Back to the proof. So *let the q we are working with* be the *smallest* that satisfies (3). Then we have the configuration (Recall Remark 4.2.9 (1))

$$\underbrace{xRz_1Rz_2Rz_3\ldots\mathbf{R}\boxed{z_i\underline{R}z_{i+1}\ldots}z_rRz_{r+1}\ldots Rz_{q-1}Ry}_{q\ Rs} \qquad (4)$$

Now the sequence (omitting x, as you note)

$$z_1, z_2, z_3\ldots z_i, z_{i+1}, \ldots z_r, z_{r+1}, \ldots, z_{q-1}, y$$

in (4) above contains $\boxed{q > n}$ members and as they all come from A, **not all are distinct**.

So let $z_i = z_r$ (the z_r could be as late in the sequence as y, i.e., equal to y).

Now omit the boxed part in (4). We obtain

$$xRz_1Rz_2Rz_3\ldots z_{i-1}\mathbf{R}z_rRz_{r+1}\ldots z_{q-1}Ry \qquad (5)$$
$$\|$$
$$z_i$$

which contains at least by **one** "R" fewer than the sequence (4) does —the entry "$z_i\underline{R}z_{i+1}$" (and everything else in the "..." part in the box) was removed. That is, (5) states

$$xR^{q'}y$$

with $q' < q$. Since the q in (3) was *smallest* $> n$, we *must have* $q' \leq n$

Pause. Why is $q' \leq n$? Because if $q' > n$ then q' satisfies (3) on p.118 *AND* is *SMALLER* than q contradicting the status of q!◀

$q' \leq n$ sends us to **Case 1** and we are done. □

4.4. Equivalence relations

Equivalence relations must be *ON some set A*, since we *require reflexivity* (definition below). They play a significant role in many branches of *mathematics* and even in *computer* science. For example, the minimisation process of finite automata (a topic that we will not cover) relies on the concept of equivalence relations and fast integer multiplication algorithms using the Fast Fourier Transform.

4.4.1 Definition. A relation R on A is an equivalence relation, provided it is all of

1. Reflexive
2. Symmetric
3. Transitive □

An equivalence relation on A has the effect, *intuitively*, of "grouping" elements that we view as *interchangeable in their roles*, or "equivalent", into so-called (see Definition 4.4.4 below) "*equivalence classes*" —kind of mathematical clubs!

4.4.2 Example. The following are equivalence relations

- $\{(1,1)\}$ on $A = \{1\}$.
- $=$ (or $\mathbf{1}_A$ or Δ_A) on A.
- Let $A = \{1,2,3\}$. Then $R = \{(1,2),(1,3),(2,3),(2,1),(3,1),(3,2),(1,1),(2,2),(3,3)\}$ is an equivalence relation on A.
- $\mathbb{N}^2$ is an equivalence relation on $\mathbb{N}$. □

Here is a longish, more sophisticated example, that is central in *number theory*. We will have another instalment of it after a few definitions and results.

4.4.3 Example. (Congruences) Fix an $m \geq 2$. We define the relation $\equiv_m$on $\mathbb{Z}$ by

$$x \equiv_m y \text{ iff } m \,|\, (x - y)$$

Recall that "$|$" is the "divides with zero remainder" relation.

> $a|b$, therefore, says that b is a multiple of a or a is a factor of b: $(\exists k) b = a \times k$.

A notation that is very widespread in the literature is to split the symbol "$\equiv_m$" into two and write

$$x \equiv y \pmod{m} \text{ instead of } x \equiv_m y$$

"$x \equiv y \pmod{m}$" and $x \equiv_m y$ are read "*x is congruent to y modulo m* (or just 'mod m')". Thus "$\equiv_m$" is the "congruence (mod m)" short symbol, while "$\equiv \ldots \pmod{m}$" is the long two-piece symbol. *We will be using the short symbol*.

We next verify the required properties for $\equiv_m$ to be an equivalence relation.

1. *Reflexivity*: Indeed, $m \,|\, (x - x)$, or $m \,|\, 0$, hence $x \equiv_m x$.

2. *Symmetry*: Clearly, if $m \,|\, (x - y)$, then $m \,|\, (y - x)$. I translate: If $x \equiv_m y$, then $y \equiv_m x$.

3. *Transitivity*: Let $m \,|\, (x-y)$ and $m \,|\, (y-z)$. The first says that, for some k, $x-y = km$. Similarly the second says, for some n, $y-z = nm$. Thus, adding these two equations I get $x - z = (k + n)m$, that is, $m \,|\, (x - z)$. I translate: If $x \equiv_m y$ and $y \equiv_m z$, then also $x \equiv_m z$. □ 

4.4.4 Definition. (Equivalence classes) Given an equivalence relation R on A. The *equivalence class* of an element $x \in A$ is $\{y \in A : xRy\}$. We use the symbol $[x]_R$, or just $[x]$ if R is understood, for the equivalence class.

 Since A is a set and $[x] \subseteq A$, each equivalence class is a set by 2.3.5. 

The symbol A/R denotes the *quotient class* of A with respect to R, that is,

$$A/R \stackrel{Def}{=} \{[x]_R : x \in A\}$$

□

Feb. 12, 2024

4.4.5 Remark. Suppose an equivalence relation R on A is given.

By reflexivity, xRx, for any x. Thus $x \in [x]_R$, hence all equivalence classes are *nonempty*.

Be careful to distinguish the brackets $\{\ldots\}$ from these $[\ldots]$.

It is NOT a priori obvious that $x \in [x]_R$ until you look at the definition 4.4.4! $[x]_R \neq \{x\}$ in general!

□

If A is a set and R is an equivalence relation on A, is the quotient *class* A/R —the standard symbol for this— a *set*?

4.4.6 Theorem. *A/R is a set for any set A and equivalence relation R on A.*

Proof. A/R just contains all the $[x]_R \subseteq A$ —recall, $[x]_R \overset{Def}{=} \{z \in A : xRz\}$.

So,

$$[x]_R \in A/R \Longrightarrow [x]_R \subseteq A \Longrightarrow [x]_R \in 2^A$$

Thus $A/R \subseteq 2^A$ and we are done by 2.3.5. □

4.4.7 Lemma. *Let P be an equivalence relation on A. Then $[x] = [y]$ iff xPy —where we have omitted the subscript $_P$ from the $[\ldots]$-notation.*

Proof. $(\rightarrow)$ part. Assume $[x] = [y]$.

By reflexivity, $y \in [y]$ (4.4.5).

The assumption then yields $y \in [x]$ and therefore xPy by 4.4.4.

($\leftarrow$) part. Assume xPy.

Let $z \in [x]$. Then xPz.

By *assumption* I also have yPx (by symmetry), thus, *transitivity* yields yPz. This says $z \in [y]$, proving

$$[x] \subseteq [y] \tag{1}$$

By symmetry of P, the "blue" assumption yields yPx and the three-line argument above yields $[y] \subseteq [x]$. *This and (1) yield* $[x] = [y]$. $\square$

4.4.8 Lemma. *Let R be an equivalence relation on A. Then*

(i) $[x] \neq \emptyset$, *for all* $x \in A$.

(ii) $[x] \cap [y] \neq \emptyset$ *implies* $[x] = [y]$, *for all* x, y *in* A.

(iii) $\bigcup_{x \in A} [x] = A$.

Note:

$$\bigcup_{x \in A} [x] \overset{Def}{=} \bigcup \Big\{ [x] : x \in A \Big\} = \bigcup A/R$$

Proof.

(i) 4.4.5.

(ii) Let $z \in [x] \cap [y]$. Then xRz and yRz, therefore xRz and zRy (the latter by *symmetry*); hence xRy (transitivity).

Thus, $[x] = [y]$ by Lemma 4.4.7.

(iii) The $\subseteq$-part is obvious from $[x] \subseteq A$.

The $\supseteq$-part follows from $\bigcup_{x \in A} \{x\} = A$ and $\{x\} \subseteq [x]$. □

The properties *(i)*–*(iii)* are characteristic of the notion of a *partition of a set*.

4.4.9 Definition. (Partitions) Let F be a family of subsets of A. It is a *partition of* A iff all of the following hold:

(i) For all $X \in F$ we have that $X \neq \emptyset$.

(ii) If $\{X, Y\} \subseteq F$ and $X \cap Y \neq \emptyset$, then $X = Y$.

(iii) $\bigcup F = A$. □

There is a natural affinity between equivalence relations and partitions on a set A. In fact,

4.4.10 Theorem. *Given a partition F on a set A. This leads to the definition of an equivalence relation P whose equivalence classes are precisely the sets —often called "**blocks**" or "**tiles**"— of the partition, that is $F = A/P$.*

Proof. First we define P:

$$xPy \overset{Def}{\text{ iff }} (\exists X \in F)\{x, y\} \subseteq X \tag{1}$$

Observe that

(i) P is *reflexive*: Take any $x \in A$. By 4.4.9(iii), there is an $X \in F$ such that $x \in X$, hence $\{x, x\} \subseteq X$. Thus xPx.

(ii) P is, trivially, *symmetric* since there is no order in $\{x, y\}$.

(iii) P is *transitive*: Indeed, let $xPyPz$. Then $\{x, y\} \subseteq X$ and $\{y, z\} \subseteq Y$ for some X, Y in F.

Thus, $y \in X \cap Y$ hence $X = Y$ by 4.4.9(ii). Hence $\{x, z\} \subseteq X$, therefore xPz.

So P is an equivalence relation. Let us compare its equivalence classes with the various $X \in F$.

Now $[x]_P$ (dropping the subscript $_P$ in the remaining proof) is

$$\{y : xPy\} \tag{2}$$

Let us compare $[x]$ with the *unique* $X \in F$ that ALSO contains x —why unique? By 4.4.9(ii). Thus,

$$y \in [x] \overset{(2)}{\Longleftrightarrow} xPy \overset{(1)}{\Longleftrightarrow} x \in X \wedge y \in X \overset{x \in X \text{ is } \mathbf{t}}{\Longleftrightarrow} y \in X$$

Thus $[x] = X$. □

4.4.11 Example. (Another look at congruences; Read Me!) Euclid's theorem for the division of integers states:

If $a \in \mathbb{Z}$ and $2 \leq m \in \mathbb{Z}$, then *there are* unique q and r such that

$$\boxed{a = mq + r \text{ and } 0 \leq r < m} \tag{1}$$

There are many proofs, but here is one: Fix a and $m \geq 2$. The set

$$T = \{x : 0 \leq x = a - mz, \text{ for some z}\}$$

is *not empty*. For example,

- if $a > 0$, then take $z = 0$ to obtain $x = a > 0$ in T.

- If $a = 0$, then take $z = 0$ to obtain $x = 0 \in T$.

- Finally, if $a < 0$, then take $z = -|a|$ [†] to obtain $x = -|a| + m|a| = |a|(m-1) > 0$ in T (since $m \geq 2$ we have $m - 1 \geq 1$).

Let then r be the *smallest* $x \geq 0$ in T.

[†] Absolute value.

The *corresponding* "z" to the *smallest* $x = r$ let us call q. So we have

$$a = mq + r, \text{ where } 0 \leq r \tag{2}$$

Can $r \geq m$? If so, then write $r = k + m$, where $k = r - m \geq 0$ and thus $k < r$. I got

$$a = m(q + 1) + k$$

As $k < r$, I have contradicted the minimality of r in (2) in the box above.

This proves that $r < m$.

We have proved *existence of at least one pair* q and r that works for (1) on p.132.

How about *uniqueness*?

Well, the worst thing that can happen is to have two representations 1). Here is another one:

$$a = mq' + r' \text{ and } 0 \leq r' < m \tag{2}$$

As both r and r' are $< m$, their "distance" (absolute difference) is also $< m$.[†]

Now, from (1) and (2) we get

$$m|q - q'| = |r - r'| \tag{3}$$

This cannot be unless $q = q'$ (in which case $r = r'$, therefore uniqueness is proved).

Wait: Why it "cannot be" if $q \neq q'$?

Because then $|q - q'| \geq 1$ thus the lhs of "=" in (3) is $\geq m$ but the rhs is $< m$.

[†]From $0 \leq r' < m$ I get $-m < r' \leq 0$. Using (1) (p.132), I get $-m < r - r' < m$. That is, $|r - r'| < m$.

We now take a deep breath!

Feb. 14, 2024

Now, back to congruences! The above was just a preamble!

Fix an $m > 1$ and consider the congruences $x \equiv_m y$. What are the equivalence classes?

Better question is what representative members are convenient to use for each such class? Given that $a \equiv_m r$ by (1) (p.132), and using Lemma 4.4.7 we have $[a]_m = [r]_m$.

r is a far better representative than a for the class $[a]_m$ as it is "normalised".

Thus, we have just m equivalence classes $[0], [1], \ldots, [m-1]$.

Wait! Are they *distinct*? Yes! Since $[i] = [j]$ is the same as $i \equiv_m j$ (4.4.7) and, since $0 < |i - j| < m$, m *cannot* divide $i - j$ with 0 remainder, we cannot have $[i] = [j]$ if $i \neq j$. □

4.4.12 Example. (A practical example) Say, I chose $m = 5$. Where does $a = -110987$ belong?

I.e., in which class out of $[0]_5$, $[1]_5$, $[2]_5$, $[3]_5$, $[4]_5$?

Well, let's do primary-school-learnt long division of $|a| = -a > 0$ divided by 5 and find quotient q and remainder r. We find, in this case, $q = 22197$ and $r = 2$. These satisfy

$$|a| = -a = 22197 \times 5 + 2$$

Thus,

$$a = -22197 \times 5 - 2 \tag{1}$$

(1) can be *rephrased* as

$$a \equiv_5 -2 \tag{2}$$

But easily we check that $-2 \equiv_5 3$ (since $3 - (-2) = 5$). Thus,

$$a \in [-2]_5 = [3]_5$$ □

4.4.13 Exercise. Can you now *easily* write the same a above as

$$a = Q \times 5 + R, \text{ with } 0 \leq R < 5?$$

Show all your work. □

4.4.1. Partial orders

Feb. 16, 2024

This subsection introduces *one of the most important kind of binary relations in set theory and mathematics in general*: The *partial order* relations.

4.4.14 Definition. (*Converse* or *Inverse* relation of $\mathbb{P}$) For any relation $\mathbb{P}$, the symbol $\mathbb{P}^{-1}$ is called the *converse* or *inverse* relation of $\mathbb{P}$ and is defined by

$$\mathbb{P}^{-1} = \{(x, y) : y\mathbb{P}x\} \tag{1}$$

$x\mathbb{P}^{-1}y$ iff $y\mathbb{P}x$ is an equivalence that says exactly what (1) does. □

4.4.15 Theorem. $\mathrm{dom}(\mathbb{P}) = \mathrm{ran}(\mathbb{P}^{-1})$ *and* $\mathrm{dom}(\mathbb{P}^{-1}) = \mathrm{ran}(\mathbb{P})$.

Proof. The *two columns* of the tables $\mathbb{P}$ and $\mathbb{P}^{-1}$ are the same, BUT swapped. Done.

Algebraically (formulaically) it is as easy:

$$\mathrm{dom}(\mathbb{P}) = \{y : (\exists x)y\mathbb{P}x\} = \{y : (\exists x)x\mathbb{P}^{-1}y\} = \mathrm{ran}(\mathbb{P}^{-1})$$

$$\mathrm{dom}(\mathbb{P}^{-1}) = \{y : (\exists x)y\mathbb{P}^{-1}x\} = \{y : (\exists x)x\mathbb{P}y\} = \mathrm{ran}(\mathbb{P}) \qquad \square$$

4.4.16 Example. If I take $\mathbb{P}$ to be "$<$" on $\mathbb{N}$, then $>=<^{-1}$ —i.e., $>$ IS the inverse of $<$— since

$$x > y \text{ iff } y < x$$ □

More notation!

4.4.17 Definition. (Important: "$(a)\mathbb{P}$" notation) For any relation $\mathbb{P}$ we write "$(a)\mathbb{P}$" to indicate the *class* —*possibly* proper— of **all** outputs of $\mathbb{P}$ *for input* a. That is,

$$(a)\mathbb{P} \stackrel{Def}{=} \{y : a\,\mathbb{P}\,y\}$$

If $(a)\mathbb{P} = \emptyset$, then we say "$\mathbb{P}$ is *undefined* at a" —that is, $a \notin \text{dom}(\mathbb{P})$.

The last "underlined" formula is read as "$\mathbb{P}$ is *undefined* at a".

If $(a)\mathbb{P} \neq \emptyset$, then $\mathbb{P}$ is "*defined*" at a —a does produce outputs!— that is, $a \in \text{dom}(\mathbb{P})$.

The blue underlined statement is read as "$\mathbb{P}$ is *defined* at a". □

4.4.18 Remark. (1) So, if R is an equivalence relation on a set A, then, using the above notation, $[x]_R = (x)R$.

(2) In analogy with the *set* $\{y : y < x\}$ over the natural numbers — that we call *the set of* $<$-**predecessors** of x— we have, in general, the *class of* $\mathbb{P}$-**predecessors** of x:

$$\{y : y\mathbb{P}x\} \qquad (\dagger)$$

Why "**predecessors**"? Well, for the natural number case above we note that $y < x$ is often read "y is before x".

(3) Note that

$$\{y : y\mathbb{P}a\} = \{y : a\mathbb{P}^{-1}y\} = (a)\mathbb{P}^{-1}$$

Thus, $\{y : y < a\} = (a) <$ □

4.4.19 Exercise. Give an example of a specific relation $\mathbb{P}$ and one specific input object (set or atom) a such that $(a)\mathbb{P}$ is *a proper class*.

□

4.4.20 Definition. (Partial Order) A relation $\mathbb{P}$ is called a *partial order* or just an *order*, iff it is *all of*

(1) *irreflexive* (i.e., $x\mathbb{P}y \to x \neq y$, for all x, y), or

(1′) Alternatively, *irreflexive* (i.e., $x\mathbb{P}x$ is false, for all x), and

(2) *transitive*.

It is emphasised that in the interest of generality —for much of this subsection (until we say otherwise)— $\mathbb{P}$ need not be a set.

Some people call this a *strict order* as it imitates the "<" on, say, the natural numbers. □

4.4.21 Remark. (1) We will *usually* use the symbol "$<$"

even in *the abstract setting*

to denote *any* unspecified *order* $\mathbb{P}$, and it will be pronounced "less than".

(2) If the order $<$ is a subclass of $\mathbb{A} \times \mathbb{A}$ —i.e., it is $<: \mathbb{A} \to \mathbb{A}$— then we say that $<$ *is an order on* $\mathbb{A}$.

(3) Clearly, for any order $<$ and any class $\mathbb{B}$, $< \cap (\mathbb{B} \times \mathbb{B})$ *is* an order on $\mathbb{B}$.

We call $< \cap (\mathbb{B} \times \mathbb{B})$ the *relational restriction of* $<$ *on* $\mathbb{B}$ and denote it by $< | \mathbb{B}$. That is, "keep ONLY the pairs whose input AND output components are in $\mathbb{B}$"

□ 

4.4.22 Exercise. How clearly? (re (3) above.) Give a simple, short proof.

Hint. $x\Big(< \cap (\mathbb{B} \times \mathbb{B})\Big) y$ iff $x < y$ and $\{x, y\} \subseteq \mathbb{B}$. □

4.4.23 Example. The standard concrete "less than", $<$, on $\mathbb{N}$ is an order, but $\leq$ is not (it is *not* irreflexive).

The "greater than" relation, $>$, on $\mathbb{N}$ is also an order, but $\geq$ is not.

In general, it is trivial to verify that "$\mathbb{P}$ is an order iff $\mathbb{P}^{-1}$" is an order. *Exercise*! □

4.4.24 Example. $\emptyset$ is an order.

Moreover for any $\mathbb{A}$, $\emptyset \subseteq \mathbb{A} \times \mathbb{A}$,

hence $\emptyset$ is also an order *on* $\mathbb{A}$ for the arbitrary $\mathbb{A}$. □

Feb. 26, 2024

4.4.25 Example. The relation $\in$ is *irreflexive* by the well known $A \notin A$, for all A.

It is *not* transitive though.

For example, $1 \in \{1\} \in \{\{1\}\}$ but $1 \notin \{\{1\}\}$.

So $\in$ is not an order. □

4.4.26 Example. Let $M = \Big\{\emptyset, \{\emptyset\}, \Big\{\emptyset, \{\emptyset\}\Big\}, \Big\{\emptyset, \{\emptyset\}, \Big\{\emptyset, \{\emptyset\}\Big\}\Big\}\Big\}$.

The relation

$$\varepsilon = \in \mid M$$

is transitive (and irreflexive), hence it is an order (*on* M). *Verify*!

□

4.4.27 Example. $\subset$ (same as $\subsetneq$) is an order, $\subseteq$—failing irreflexivity—is not. □

4.4.28 Example. (Why "Partial" Order?) Consider the order $\subset$ again.

In this case we have **all of** $\{1\} \subset \{\{1\}\}$, $\{\{1\}\} \subset \{1\}$ or $\{\{1\}\} = \{1\}$ **false**.

That is, $\{1\}$ and $\{\{1\}\}$ are *non comparable* items.

This justifies the qualification *partial* for orders in general (Definition 4.4.34).

On the other hand, the "natural" $<$ on $\mathbb{N}$ is such that one of $x = y$, $x < y$, $y < x$ always holds for any x, y in $\mathbb{N}$.

That is, all (unordered) pairs x, y of $\mathbb{N}$ *are* comparable under $<$.

While *all* orders are "partial", some are *total* ($<$ above) and others are *nontotal* ($\subset$ above).

"Partial" is *not* the negation of "total".

□ 

4.4.29 Definition. Let $<$ be an **arbitrary** (abstract) partial order on $\mathbb{A}$. We define

$$\leq \stackrel{Def}{=} \mathbf{\Delta}_{\mathbb{A}} \cup <$$

We pronounce $\leq$ "less than or equal".

$\mathbf{\Delta}_{\mathbb{A}} \cup >$ is denoted by $\geq$ and is pronounced "greater than or equal".

Let us call $\leq$ a *reflexive order* or also a *non strict order*.

□

(1) In plain English, given $<$ on $\mathbb{A}$, we *have defined* $x \leq y$ to mean

$$x < y \vee \overbrace{x = y}^{\text{equality is } \Delta_{\mathbb{A}}}$$

for all x, y in $\mathbb{A}$.

(2) The definition of $\leq$ *depends* on the *FIELD* $\mathbb{A}$ due to the presence of $\mathbf{\Delta}_{\mathbb{A}}$.

There is no need for such dependency on any "reference" class (*Field*) in the case of $<$.

Feb. 28, 2024

Recall that "$<$" —as in lemma below— will be used often, with warning, as an "abstract" (= unspecified) order other than the familiar one on $\mathbb{N}$ or $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$.

4.4.30 Lemma. *For any abstract —that is, not specific— $<: \mathbb{A} \to \mathbb{A}$, the associated relation $\leq$ on $\mathbb{A}$ defined in 4.4.29 is reflexive, antisymmetric and transitive.*

Proof.

(1) *Reflexivity* is trivial. $\Delta_{\mathbb{A}}$ "throws in" all pairs (x, x) for all $x \in \mathbb{A}$.

(2) For *antisymmetry*, **let** $\underbrace{x \leq y}_{x=y \vee x<y}$ and $\underbrace{y \leq x}_{x=y \vee y<x}$.

I will prove $x = y$ **by contradiction**.

Suppose $x \neq y$ instead. Then hypothesis (the "**let**"-sentence above) becomes $x < y$ and $y < x$, hence (by transitivity of "$<$") $x < x$. This contradicts *irreflexivity* of $<$.

(3) As for *transitivity* **Let** $x \leq y$ and $y \leq z$.

We want to prove that $x \leq z$ **follows from hypothesis (3)**.

(a) If $x = z$ **we are done**, since then $x \leq z$ is true: $\overbrace{\underbrace{x = z}_{\mathbf{t}} \vee x < z}^{\mathbf{t}}$.

(b) **The remaining case is** $x \neq z$

The Subcases below analyse hypothesis (3) —the "Let"-sentence above.

- Subcase $x = y$. Then $y \leq z$ (see (3)) becomes $x \leq z$. Done.
- Subcase $y = z$. Then $x \leq y$ (see (3)) becomes $x \leq z$. Done.
- Subcase $x \neq y$ AND $y \neq z$ (the subcase $x = y = z$ is **impossible** given that $x \neq z$).

  So we have (by (3)) $x < y$ and $y < z$

  By transitivity of $<$ we get $x < z$, hence $x \leq z$, since the latter says $\underbrace{x < z}_{\mathbf{t}} \vee x = z$. **Done one last time!** □

4.4.31 Lemma. *Let $\mathbb{P}$ on $\mathbb{A}$ be reflexive, antisymmetric and transitive. Then $\mathbb{P} - \mathbf{\Delta}_{\mathbb{A}}$ is a (strict) order on $\mathbb{A}$.*

Proof. Since

$$\mathbb{P} - \mathbf{\Delta}_{\mathbb{A}} \subseteq \mathbb{P} \tag{1}$$

it is clear that $\mathbb{P} - \mathbf{\Delta}_{\mathbb{A}}$ is *on* $\mathbb{A}$.

It is also clear that it is *irreflexive* since we REMOVED ALL (x, x) pairs, which are in $\mathbf{\Delta}_{\mathbb{A}}$. We only need verify that it is *transitive*.

So let

$$(x, y) \text{ and } (y, z) \text{ be in } \mathbb{P} - \mathbf{\Delta}_{\mathbb{A}} \tag{2}$$

We want $(x, z) \in \mathbb{P} - \mathbf{\Delta}_{\mathbb{A}}$

By (1) and (2)

$$(x, y) \text{ and } (y, z) \text{ are in } \mathbb{P} \tag{3}$$

hence

$$(x, z) \in \mathbb{P} \tag{4}$$

by the given *transitivity* of $\mathbb{P}$.

$$\text{But I want } (x, z) \in \mathbb{P} - \Delta_{\mathbb{A}} \tag{†}$$

Can $(x, z) \in \mathbf{\Delta}_{\mathbb{A}}$, i.e., can $x = z$?

No, because antisymmetry of $\mathbb{P}$ (given) and (3) would then imply $x = y$, i.e., $(x, y) \in \mathbf{\Delta}_{\mathbb{A}}$ *contrary* to (2).

So, $(x, z) \in \mathbb{P} - \mathbf{\Delta}_{\mathbb{A}}$ by (4), and we got (†). □

4.4.32 Corollary. *Let $\leq$ on $\mathbb{A}$ be reflexive, antisymmetric and transitive. Then $<$ defined by*

$$x < y \overset{Def}{\equiv} x \leq y \wedge x \neq y$$

is a (strict) order on $\mathbb{A}$.

Proof. The corollary just rephrases 4.4.31 in a different notation. □

4.4.33 Remark. Lemmas 4.4.30 and 4.4.31 show that the two approaches —"$<$" and "$\leq$"— are interchangeable. However the "modern" approach of Definition 4.4.20 avoids the nuisance of having to tie the notion of order to some particular "field" $\mathbb{A}$ (4.1.7).

For us, in class and in our notes, "$\leq$" is the *derived, secondary* notion defined in 4.4.29. □

4.4.34 Definition. (PO Class) If $<$ is an order *on* a class $\mathbb{A}$, we call the *informal* **pair** $(\mathbb{A}, <)$ a *partially ordered class*, or *PO class*.

If $<$ is an order on a *set* A, we call the pair $(A, <)$ a *partially ordered set* or *PO set*. Often, if the order $<$ is understood as being on $\mathbb{A}$ or A, one says that "$\mathbb{A}$ is a PO class" or "A is a PO set" respectively. □

Mathematically speaking, $(\mathbb{A}, <)$ is *not* an ordered pair when $\mathbb{A}$ is a *proper* class because in $\{\mathbb{A}, \{\mathbb{A}, <\}\}$ we do not allow class *members*. We may think instead (non mathematically) of "$(\mathbb{A}, <)$" as *informal* notation that simply "associates" $\mathbb{A}$ and $<$ together into a "toolbox" $(\ldots, \ldots)$.

4.4.35 Definition. (Linear order) A relation $<$ on $\mathbb{A}$ is a *total* or *linear* order *on* $\mathbb{A}$ iff it is all of

(1) An order, and

(2) For any x, y in $\mathbb{A}$ **one of** $x = y, \quad x < y, \quad y < x$ **is true** —this is the so-called "*trichotomy*" property.

Trichotomy says: For any x, y we have $x = y \vee x < y \vee x > y$ is true

If $\mathbb{A}$ is a class, then the informal pair $(\mathbb{A}, <)$ is a *linearly ordered class* —in short, a *LO class*.

If $\mathbb{A}$ is a set, then the pair $(\mathbb{A}, <)$ is a *linearly ordered set* —in short, a *LO set*.

One often calls just $\mathbb{A}$ a LO class or LO set (as the case warrants) when $<$ is understood from the context. □

4.4.36 Example. The standard $<: \mathbb{N} \to \mathbb{N}$ is a total order, hence $(\mathbb{N}, <)$ is a LO set.

Mar. 1, 2024

4.4.37 Definition. (Minimal and minimum elements) Let $<$ be **ANY** (irreflexive) order and $\mathbb{A}$ be any class.

We are *NEITHER* requiring NOR assuming that $<$ is *ON* $\mathbb{A}$.

An element $b \in \mathbb{A}$ is *a <-**minimal** element IN $\mathbb{A}$*, or *a <-minimal element of $\mathbb{A}$*, or *minimal in $\mathbb{A}$* with respect to $<$, iff

$$\neg(\exists x \in \mathbb{A})x < b$$

or

$$\mathbb{A} \cap \{x : x < b\} = \emptyset$$

In words, there is nothing before b in $\mathbb{A}$.
b has NO "predecessors" (see Remark 4.4.18, item (2)) in $\mathbb{A}$.

$m \in \mathbb{A}$ is *a <-**minimum** element IN $\mathbb{A}$* iff $(\forall x \in \mathbb{A})m \leq x$.[†]

If $a \in \mathbb{A}$ is $>$-minimal in $\mathbb{A}$, that is $\neg(\exists x \in \mathbb{A})x > a$, we call a a *$<$-maximal* element in $\mathbb{A}$. Similarly, a $>$-minimum element is called a *$<$-maximum*.

If $<$ is understood, then the qualification "<-" is omitted. □

[†]Of course, "$m \leq x$" says (means) $m < x \vee m = x$.

4.4.38 Exercise. In particular, if b ($\in \mathbb{A}$) is *not* in the *field*

$$\mathrm{dom}(<) \cup \mathrm{ran}(<)$$

(cf. 4.1.7) of $<$, then b is $<$-minimal *in* $\mathbb{A}$.

Hint. Compute $\{x : x < b\}$. □

4.4.39 Remark. (Important) Note how the notation learnt from 4.4.17 can *simplify* the expression

$$\neg(\exists x \in \mathbb{A})x < a \tag{1}$$

Since $x < a$ iff $a > x$, (1) says that *no* x *is in* ***both*** $\mathbb{A}$ and in the *predecessor class* $\{x : x < a\} = \{x : a > x\} = (a) >$.[†]

That is, a is $<$-minimal in $\mathbb{A}$ iff

$$\boxed{\mathbb{A} \cap (a) > = \emptyset} \tag{2}$$

□

[†] $\underbrace{\{x : x < a\}}_{\text{class of predecessors of } a} = \{x : a > x\} = (a) >$ (see also 4.4.18).

4.4.40 Example. (Important) 0 is *minimal*, also *minimum*, in $\mathbb{N}$ with respect to the natural ordering.

In $\mathscr{P}(\mathbb{N})$, $\emptyset$ is both $\subset$-minimal and $\subset$-minimum.

On the other hand, all of $\{0\}, \{1\}, \{2\}$ are $\subset$-minimal in $\mathscr{P}(\mathbb{N}) - \{\emptyset\}$ but *none* are $\subset$-*minimum* in that set. For example, $\{1\} \not\subseteq \{2, 3\}$.

So, the concepts "minimal" and "minimum" are DISTINCT!

Observe from this last example that minimal elements in a class are *not* unique.

□ 

4.4.41 Remark. (Hasse diagrams) Read me! There is a neat pictorial way to depict orders on finite sets known as "*Hasse diagrams*". To do so one creates a so-called "*graph*" of the finite PO set $(A, <)$ where $A = \{a_1, a_2, \ldots, a_n\}$.

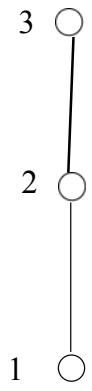
How? The graph consists of n *nodes* —which are drawn as points— each labeled by one a_i. The graph also contains 0 or more *arrows* that connect nodes. These arrows are called *edges*.

When we depict an arbitrary R on a finite set like A we draw *one* arrow (edge) from a_i to a_j iff the two *relate*: $a_i R a_j$.

In Hasse diagrams for PO sets $(A, <)$ we are more selective: We say that b *covers* a iff $a < b$, but there is no c such that $a < c < b$. In a Hasse diagram we will

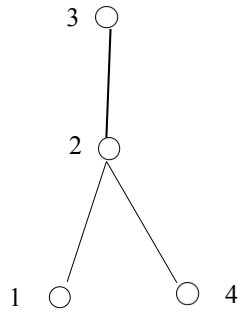
1. draw an edge from a_i to a_j **iff** a_j covers a_i.
2. by convention we will draw b **higher** than a on the page if b covers a.
3. given the convention above, **using "arrow-heads" is superfluous**: our edges are plain line segments.

So, let us have $A = \{1, 2, 3\}$ and $<= \{(1, 2), (1, 3), (2, 3)\}$.



The above has a minimum (1) and a maximum (3) and is clearly a linear order.

A slightly more complex one is this $(A, <)$, where $A = \{1, 2, 3, 4\}$ and $<= \{(1, 2), (4, 2), (2, 3), (1, 3), (4, 3)\}$.



This one has a maximum (3), two minimal elements (1 and 4) but no minimum, and is not a linear order: 1 and 4 are not comparable. □

4.4.42 Lemma. *Given an order $<$ and a class $\mathbb{A}$.*
(1) If m is a minimum *in $\mathbb{A}$, then it is also* minimal.
(2) If m is a minimum *in $\mathbb{A}$, then it is unique.*

Proof. (1) Let m be mini*mum* in $\mathbb{A}$. Then

$$m \leq x, \text{ that is, we have } m = x \vee m < x \qquad (i)$$

for all $x \in \mathbb{A}$.

Now, prove that there is no $x \in \mathbb{A}$ such that $x < m$.

OK, let us go *by contradiction*:

- So let instead, for some $a \in \mathbb{A}$,

$$a < m \qquad (ii)$$

that is, suppose m is NOT minimal.

- I also have $m \leq a$ by (i), because both m and a are in $\mathbb{A}$ and m is *minimum*; that is,

$$\overbrace{m = a}^{\text{f by } (ii)} \vee m < a \qquad (iii)$$

- So, (iii) *nets* $m < a$.

So (ii) and (iii) and transitivity yield $a < a$; contradiction ($<$ is irreflexive). Done.

(2) Let m and n both be *minima* (*plural of minimum*) in $\mathbb{A}$. Then $m \leq n$ (with m posing as minimum) and $n \leq m$ (now n is so posing), hence $m = n$ by antisymmetry (Lemma 4.4.30). □

4.4.43 Lemma. *If $<$ is a* linear *order on $\mathbb{A}$, then every minimal element is also minimum.*

Proof. Easy **Exercise**!

Hint. If $a \in \mathbb{A}$ is minimal, then, for all $x \in \mathbb{A}$, the statement "$x < a$" is false. Since for all x the statement $x < a \vee a < x \vee x = a$ is true (because $<$ is total), we have for all x the statement $a < x \vee x = a$ is true. ETC. □

So, by 4.4.42 and 4.4.43,

4.4.44 Corollary. *In a linear order the concepts miniMUM and miniMAL coincide.*

The following type of relation has fundamental importance for set theory, and mathematics in general.

4.4.45 Definition.

1. A general (irreflexive) order $<$ satisfies the *minimal condition*, in short *it has MC*, iff *EVERY* nonempty $\mathbb{A}$ "out there"[†] *DOES have* $<$-minimal elements.

2. If a *total* order $<: \mathbb{B} \to \mathbb{B}$ has MC, then it is called a *well-ordering*[‡] *on* (or *of*) the class $\mathbb{B}$.

3. If $(\mathbb{B}, <)$ is a LO class (or LO set) where "$<$" has MC, then it is a *well-ordered class* (or well-ordered set), or *WO class* (or WO set).

□

[†]This "out there" implies that $\mathbb{A}$ is not in any way tied or connected to $<$ (as a field or whatever).
[‡]The term "well-ordering" is ungrammatical, but it is *the* terminology established in the literature!

Mar. 6, 2024

4.4.46 Remark.

In symbols, Definition 4.4.45, **Item 1**, says that $<$ has MC iff the following is true:

$$\emptyset \neq \mathbb{A} \to (\exists a \in \mathbb{A}) \overbrace{\underbrace{\mathbb{A} \cap (a) >= \emptyset}_{a \text{ is } <\text{-minimal in } \mathbb{A}}}^{\neg(\exists x \in \mathbb{A}) x<a} \tag{1}$$

The following **REPHRASING of (1)** is very important *for future reference*:

If $\mathbb{A}$ is given via a defining property $F(x)$, as $\mathbb{A} =^{Def} \{x : F(x)\}$, then (1) translates —in terms of $F(x)$— into

$$\overbrace{(\exists a) F(a)}^{\mathbb{A} \neq \emptyset} \to \overbrace{(\exists a)}^{(\exists a \in \mathbb{A})} \Big(F(a) \wedge \overbrace{\neg \underbrace{(\exists y) \big(F(y)}_{(\exists y \in \mathbb{A})}}^{\neg(\exists y \in \mathbb{A})} \wedge a > y) \Big) \tag{2'}$$

OR

$$(\exists a) F(a) \to (\exists a) \Big(F(a) \wedge \neg (\exists y) \big(y < a \wedge F(y) \big) \Big) \tag{2}$$

Chapter 5

Functions

We consider here a *special case of relations* that we know as "functions".

Many of you know already that a function is a relation with some special properties.

Let's make all this official:

5.1. Preliminaries

5.1.1 Definition. A *function* $\mathbb{R}$ is a *single-valued* relation.

That is,

whenever we have *both* $x\mathbb{R}y$ and $x\mathbb{R}z$

then

we will also have $y = z$ □

NOTATION. It is traditional to use, generically, lower case letters from among f, g, h, k when dealing with functions that are sets and $\mathbb{F}, \mathbb{G}, \mathbb{H}, \mathbb{K}$ for functions that are proper classes —with primes and/or subscripts if we run out of letters.

The above definition of "function" does not care about *left* or *right fields*.

5.1.2 Remark. Another way of putting it, using the notation from 4.4.17, is:

> A relation $\mathbb{R}$ is a function *iff*, for each a, $(a)\mathbb{R}$ is either *empty* or a **singleton** (i.e., contains *exactly one* element).

□ 

5.1.3 Example. (Important) The empty set is a relation of course, the empty set of *pairs*. *It is also a function since*

$$\overbrace{(x,y) \in \emptyset \wedge (x,z) \in \emptyset}^{\mathbf{f}} \rightarrow y = z$$

vacuously, by virtue of the left hand side of $\rightarrow$ being false. □

5.1.4 Example. (Important) The diagonal $\mathbf{1}_{\mathbb{A}} : \mathbb{A} \rightarrow \mathbb{A}$ is a function. Indeed,

$$\text{For any } x \in \mathbb{A} \text{ we have } (x)\Delta_{\mathbb{A}} = \{x\}$$

□

5.1.5 Definition. (Function-specific notations and concepts)
Let $\mathbb{F}$ be a function.

1. First off, the *concepts* *AND* *notation* for
   - domain
   - range,

   and —*in case of* a function $\mathbb{F} : \mathbb{A} \to \mathbb{B}$
   - left field
   - right field
   - field
   - total

   and
   - onto

   are inherited from those for relations without change.

2. Even the notations "$a\mathbb{R}b$", "$(a, b) \in \mathbb{R}$" and "$(a)\mathbb{R}$" transfer over to functions and are OFTEN useful and *ARE* employed!

3. And yet, we have an annoying *difference* in notation:

For a relation $\mathbb{F}$ —or viewing a function $\mathbb{F}$ as a relation— the class

$$\{y : a\mathbb{F}y\} \tag{1}$$

is denoted by $(a)\mathbb{F}$ (first defined in 4.4.17).

If $\mathbb{F}$ is a function, then the class in (1) is either *empty* or has ONE element ONLY (see 5.1.2); say, y.

In Relational Notation that is:

$$(a)\mathbb{F} = \begin{cases} \{y\} & \text{if } \mathbb{F} \text{ defined at } a \\ \emptyset & \text{if } \mathbb{F} \text{ undefined at } a \end{cases} \tag{2}$$

The *literature* in general[b] denotes (2) in this "function-specific" NOTATION

$\mathbb{F}(a)= y$ $\quad \langle$note order reversal from $(a)\mathbb{F}$ and braces-removal!$\rangle$

$\mathbb{F}(a)\uparrow$ $\quad \langle \mathbb{F}$ undefined at $a\rangle$

[b] Not all the literature: The significant book [Kur63] writes "af" for (set) functions AND relations, omitting even the brackets around a.

NOTATION: Thus for a *function* $\mathbb{F}$, we have **all the notations below available** to us!

$$a\mathbb{F}y \text{ iff } (a)\mathbb{F} = \{y\} \text{ iff } \boxed{\mathbb{F}(a) = y}$$

and

$$\neg(\exists y)a\mathbb{F}y \text{ iff } (a)\mathbb{F} = \emptyset \text{ iff } \boxed{\mathbb{F}(a)\uparrow}$$

□

5.1.6 Example. (Read Me!) In particular $\mathbb{F}(a) = \emptyset$ means

$$(a)\mathbb{F} = \{\emptyset\}$$

that is, $(a, \emptyset) \in \mathbb{F}$ or $a\mathbb{F}\emptyset$ —not what one might hastily think it means!

Definitely, $\mathbb{F}(a) \downarrow$ here, with output the object "$\emptyset$", it is NOT $\mathbb{F}(a) \uparrow$

□ 

5.1.7 Definition. (Images) The class of *all* outputs of a function $\mathbb{F}$, *when* **all the inputs come from any particular class** $\mathbb{X}$, is called the *image of* $\mathbb{X}$ *under* $\mathbb{F}$ and is denoted by $\mathbb{F}[\mathbb{X}]$.

Thus, mathematically,

$$\mathbb{F}[\mathbb{X}] \overset{Def}{=} \{ \overbrace{\mathbb{F}(x)}^{all\ outputs\ for\ x\in\mathbb{X}} : x \in \mathbb{X}\} \tag{1}$$

Note that careless notation like $\mathbb{F}(A)$ —where A is a set— will *not* do for $\mathbb{F}[A]$.

The ( )-notation means the input *IS* **THE** object A —*NOT* **members** of A.

If I want the inputs to be *FROM* **INSIDE** A, then I *MUST use* [ ]*-notation; I did*!

The *inverse image* of a class $\mathbb{Y}$ under a function $\mathbb{F}$ is useful as well, that is, the class of *all* inputs that **cause** $\mathbb{F}$-outputs exclusively in $\mathbb{Y}$.

It is denoted by $\mathbb{F}^{-1}[\mathbb{Y}]$ and is defined as

$$\mathbb{F}^{-1}[\mathbb{Y}] \stackrel{Def}{=} \{x : \mathbb{F}(x) \in \mathbb{Y}\} \tag{2}$$

There may well exist $y \in \mathbb{Y}$ such that NO x exists such that $\mathbb{F}(x) = y$.

For example if $\mathbb{F} = \{(0, 1)\}$ and $\mathbb{Y} = \{3\}$, then $\mathbb{F}^{-1}[\mathbb{Y}] = \emptyset$. **No input causes output 3.**

□

This is a good time to introduce "**Principle 3**"[†] of set formation.

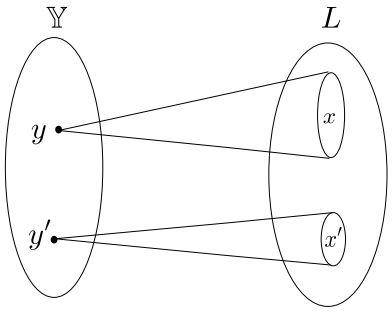
5.1.8 Remark. (LABELLING) "**Suppose that the *class* (of sets and/or atoms) $\mathbb{Y}$ *is indexed/labelled* by some (or all) members of a *set* L. Then $\mathbb{Y}$ too is a set".**

I am using "**INDEXED**" as synonymous to "**LABELLED**" by (some) members of a *set* L so that, to every $X \in \mathbb{Y}$, we have attached as "*LABEL(S)*" OR "*INDICES*" (often in form of subscripts or superscripts) some member(s) of L.

REQUIREMENT on LABELS: *We may label any member of* $\mathbb{Y}$ *with many labels* from L, but we *may NEVER use* **the same label twice** for labelling, and *may NOT leave any member of* $\mathbb{Y}$ *unlabelled*.

Example. If $\mathbb{Y} = \{A, B, C\}$, then $\{A_1, B_{13,19,0}, C_{42}\}$ is a valid labelling with labels **from** $\mathbb{N}$.[‡]

$\{A_{1,13}, B_{13}, C_{19}\}$ is not correctly labelled (same label twice), the labelling of $\{A_{1,42}, B_{13}, C\}$ is also invalid (C was not labelled):



[†]This is the last Principle; I promise!
[‡]B has three labels attached to it.

Note that $\mathbb{Y}$, intuitively speaking, has *no "MORE" members* than the label *set* L since for EACH **ONE** member $A \in \mathbb{Y}$, we SPEND **ONE** or **MORE** *labels* from the label *set* L, and none of these labels REPEATS. See preceding figure.

Thus our *intuition can accept that* $\mathbb{Y}$ *is not "bigger" than the label set*, L.

This intuitive acceptance is made "Official" via

PRINCIPLE 3: A class $\mathbb{Y}$ is proved to be a *set* as long as it has a labelling with labels from a *set* L.

Some people call Principle 3 the **"size limitation doctrine"**.

Researchers on the foundations of set theory felt that paradoxes occurred in connection with "enormous classes".

Why? Because, intuitively, when building "enormous classes" **we run out of stages needed to build them as** *SETS*.

So "small" is good and Principle 3 helps us discover NEW "small" classes (therefore *sets*) **by comparing them** with known to us "small" *label-classes*. □

5.1.9 Theorem. *If $\mathbb{G}$ is a function, and L is a set, then $\mathbb{G}[L]$ is a set.*

Proof. Let

$$\mathbb{Y} = \mathbb{G}[L] \tag{†}$$

See figure on p.181.

This means that, for every y, $y \in \mathbb{Y}$ iff for some $x \in L$, $\mathbb{G}(\mathbf{x}) = \mathbf{y}$.

The box above says the same thing as

EACH $y \in \mathbb{Y}$ is *LABELLED* by *ALL* the x that satisfy

$$x \in L \text{ and } \overbrace{\underbrace{\mathbb{G}(x) = y}_{OR,\ same\ as\ x \in \mathbb{G}^{-1}[\{y\}]}}^{same\ as\ \ \mathbb{G}(x) \in \{y\}} \tag{‡}$$

By (‡), the labels (**from** L) used for EACH y are **PRECISELY** all the x satisfying

$$x \in L \cap \mathbb{G}^{-1}[\{y\}] \text{ true because the Box is} \tag{¶}$$

Refer again to the figure on p.181

- All members of $\mathbb{Y}$ do receive labels from L by the non-emptiness of the set in (¶) [Why "set"?]
- If $y \neq y'$, both in $\mathbb{Y}$, then they receive non overlapping labels from $L \cap \mathbb{G}^{-1}[\{y\}]$ and $L \cap \mathbb{G}^{-1}[\{y'\}]$ —as in the preceding drawing of p.181— because $(L \cap \mathbb{G}^{-1}[\{y\}]) \cap (L \cap \mathbb{G}^{-1}[\{y'\}]) = \emptyset$ since

  $$\mathbb{G}^{-1}[\{y\}] \cap \mathbb{G}^{-1}[\{y'\}] = \emptyset$$

  Indeed, if $z \in \mathbb{G}^{-1}[\{y\}] \cap \mathbb{G}^{-1}[\{y'\}]$, then $\mathbb{G}(z) = y$ *and* $\mathbb{G}(z) = y'$; impossible for a function.

By **Principle 3**, $\mathbb{Y}$ —being labelled by members of the *set* L— is a set too. □

5.1.10 Corollary. *If* $\mathbb{G}$ *is a function and* $\mathrm{dom}(\mathbb{G})$ *is a set, then* $\mathbb{G}$ *is a set.*

Proof. *Exercise*! □

Pause. So far we have been giving definitions regarding functions of *one* variable. Or have we?◀

Not really: We have already said that the multiple-input case is subsumed by our notation. If $\mathbb{F} : \mathbb{A} \to \mathbb{B}$ and $\mathbb{A}$ is a class of n-tuples, then $\mathbb{F}$ is a function of "n-variables".

The binary relation, that such an $\mathbb{F}$ is, contains pairs like $\big((\vec{x}_n), x_{n+1}\big)$.

However, we usually abuse the notation $\mathbb{F}\Big((\vec{x}_n)\Big)$ —or $\Big((\vec{x}_n)\Big)\mathbb{F}$— and write instead $\mathbb{F}(\vec{x}_n)$ —or $(\vec{x}_n)\mathbb{F}$— omitting *the brackets of the n-tuple* $(\vec{x}_n)$.

5.1.11 Remark. (READ ME!) Regarding, say, the definition of $\mathbb{F}[X]$ (5.1.7):

What if $\mathbb{F}(a)\uparrow$*? How do you "collect" an* undefined "value" *into a class?*

Well, you don't.

Both (1) and (2) in 5.1.7 have a rendering that is *independent* of the notation "$\mathbb{F}(a)$" or even "$(a)\mathbb{F}$".

$$\mathbb{F}[\mathbb{X}] = \{y : (\exists x \in \mathbb{X})x\mathbb{F}y\} \tag{1'}$$

$$\mathbb{F}^{-1}[\mathbb{Y}] = \{x : (\exists y \in \mathbb{Y})x\mathbb{F}y\} \tag{2'}$$

□ 

5.1.12 Example. (Important) Thus, $f[\{a\}] = \{f(x) : x \in \{a\}\} = \{f(x) : x = a\} = \{f(a)\}$.

Let now $g = \Big\{(1,2), \Big(\{1,2\},2\Big), (2,7)\Big\}$, clearly a function. Thus, $g(\{1,2\}) = 2$, but $g[\{1,2\}] = \{2,7\}$. Also, $g(5) \uparrow$ and thus $g[\{5\}] = \emptyset$.

On the other hand, $g^{-1}[\{2,7\}] = \{1,\{1,2\},2\}$ and $g^{-1}[\{2\}] = \{1,\{1,2\}\}$, while $g^{-1}[\{8\}] = \emptyset$ since no input causes output 8. □

5.1.13 Remark. (Kleene Equality) When $f(a)\downarrow$, then $f(a) = f(a)$ as is naturally expected.

What about when $f(a)\uparrow$?

This begs a more general question that we settle as follows (following Kleene, [Kle43]):

When is $f(a) = g(b)$ where f, g are two functions?

Intuitive answer: $f(a) = g(b)$ IFF the two function "calls" left and right of "=" **produce the SAME RESPONSE**.

In symbols:

$$f(a) = g(b) \overset{Def\ ([\text{Kle43}])}{\equiv} f(a)\uparrow \wedge\, g(b)\uparrow \vee (\exists y)\Big(f(a) = y \wedge\ g(b) = y\Big)$$

□

Mar. 8, 2024

5.1.14 Example. Let $g = \{(1,2), (\{1,2\}, 2), (2,7)\}$.

Then, $g(1) = g(\{1,2\})$ and $g(1) \neq g(2)$.

$g(3) = g(4)$ since both sides are undefined. □

5.1.15 Definition. A function f is 1-1 **iff** (i.e., the concept "**1-1**" **is short for) for all** x, y **and** z**,** $f(x) = f(y) = z$ **implies** $x = y$**.**

This means the SAME, in relational notation, AS:

$$f \text{ is 1-1 iff } xfz \wedge yfz \to x = y \tag{1}$$

In words, the above says: distinct inputs must cause distinct outputs.

Same definition for a possibly non-set function $\mathbb{F}$. □

Wait! Why does our definition say distinct inputs **"map" to** (= "**produce**") *distinct results*?

Well take the **contrapositive** of (1):

For two statements S and S', the **contrapositive** of the implication $S \to S'$ is $\neg S' \to \neg S$.

$$\overbrace{x \neq y}^{\text{suppose } \mathbf{t}} \to \neg\Big(\overbrace{xfz}^{\text{suppose } \mathbf{t}} \wedge \overbrace{yfz}^{\text{must be } \mathbf{f}} \Big)$$

That is, if the inputs are different and one (the x) produces z, then the other (the y) *cannot also* produce z.

5.1.16 Remark. You might ask, "What's wrong with defining f is 1-1 by simply requiring $f(x) = f(y) \to x = y$? I saw this in dubious texts."

1-1-ness is RELEVANT to ANY function, total or not. However, dubious texts believe all functions are total. For example the function $f = \{(1,2),(2,9),(3,8)\}$ is 1-1 according to intuitive expectations that are respected by **the correct** definition:

Distinct inputs $1, 2, 3$ produce distinct actual outputs $2, 9, 8$.

If we used the dubious (and wrong) definition (plenty of "lost" discrete "MATH" books out there!) this f *would not be 1-1* since, for example, we have $f(4)\uparrow = f(5)\uparrow$, yet $4 \neq 5$.

Our definition supports what we immediately see: f *IS* 1-1. □

5.1.17 Example. (Important) $\{(1,1)\}$ and $\{(1,1),(2,7)\}$ are 1-1: Also,

$\emptyset$ is 1-1 *vacuously*.

$\{(1,0),(2,0)\}$ is *not* 1-1. □

5.1.18 Exercise. (Important) Prove that if f is a 1-1 function, then the *relation converse* f^{-1} is a function (that is, a single-valued relation). □

5.1.19 Definition. (1-1 Correspondence) A function $f : A \to B$ is called a *1-1 correspondence* iff it is all three: 1-1, total, and onto.

Often we say that "A and B are *in 1-1 correspondence*" writing

$$A \overset{f}{\sim} B$$

often omitting mention of the function that *is* the 1-1 correspondence. □

5.1.20 Exercise. Show that $\sim$ is a *symmetric* and *transitive* relation on sets. □

5.1.21 Remark. (Composition Again!) The concept of composition is *NOT NEW*. Functions *ARE* relations, so we know what composition is!

Thus, $f \circ g$ for two functions still means

$$x\, f \circ g\, y \text{ iff, for some } z,\ x\, f\, z\, g\, y \tag{1}$$

▶ **But also Note!**

$f \circ g$ is also a function. Indeed, if we have

$$x f \circ g\, y \text{ and } x f \circ g\, y'$$

then

$$\text{for some } z, x f z g\, y \tag{2}$$

and

$$\text{for some } w, x f w g\, y' \tag{3}$$

Since f is a function, (2) and (3) give $z = w$. In turn, this (since g is a function too!) gives $y = y'$. □

The notation (as in 4.4.17) "$(a)f$" for relations is "uncommon"[†] when applied to functions —but it IS correct— where "$f(a)$" may be more convenient and more "usual".

However, the "function" notation "$f(a)$" is awkward in connection with composition.

If we write $(f \circ g)(a)$ this *might* be misread as if g grabs the input! But it is f that "acts first".

We want the action $g\Big(f(a)\Big)$.

[†]See however [Kur63].

We need a new notation (below) for functional composition.

5.1.22 Definition. (Salvaging Notation "$f(a)$")
The present definition is *about NOTATION only*.

Let f and g be two functions. The Notation $f \circ g$, their *relational composition*, is the one in 4.2.1.

However, for composition of *functions*, we *ALSO* have the alternative *functional notation for composition*:

"gf" stands for "$f \circ g$"; *note the order reversal* **AND** the **absence** of "$\circ$", the composition symbol.

In particular we *write* $(gf)(a)$ *for* $(a)(f \circ g)$ —cf. 5.1.5— placing the input close to the function that uses it.

Thus let f and g be functions, hence as we saw (5.1.21), $f \circ g$ is a function as well.

Therefore

$$\begin{aligned}
\Big(gf\Big)(a) = b \text{ iff } & (a)(f \circ g) = \{b\} \text{ (\textbf{Box} on p.197 via the lens of p.177)} \\
& \text{iff } a(f \circ g)b \\
& \text{iff } (a)f = \{c\} \wedge (c)g = \{b\}, \text{ for some } c \\
& \text{iff } f(a) = c \wedge g(c) = b, \text{ for some } c \\
& \text{iff } g\Big(f(a)\Big) = b
\end{aligned}$$

The two *reds* in the formula display above uphold the intuition that f gets its input first and passes its output as input to g. □

5.1.23 Theorem. *Functional composition is associative, that is,*

$$(gf)h = g(fh)$$

Proof. Exercise!

Hint. Note that by, 5.1.22, $(gf)h = h \circ (f \circ g)$. Take it from here. □

5.1.24 Example. (Important! We know this from 5.1.4)

The *identity relation* on a set A is a function since $(a)\mathbf{1}_A$ is the *singleton* —meaning "one-element" set— $\{a\}$.

In functional notation, $\mathbf{1}_A(a) = a$ □

The following interesting result connects the notions of ontoness and 1-1ness with the "algebra" of composition.

5.1.25 Theorem. *Let $f : A \to B$ and $g : B \to A$ be functions. If*

$$gf = \mathbf{1}_A \tag{1}$$

then g is **onto** *while f is total and* **1-1**.

5.1.26 Definition. Relating to (1) in the theorem above we say that g is a **left inverse** of f and f is a **right inverse** of g.

Using the indefinite article "a" because these are not in general unique! Read Examples 5.1.27 and 5.1.28!

□

Proof. (**of 5.1.25**)

About g: Our goal, *ontoness*, means that, for each $x \in A$, I can "solve the equation $g(y) = x$ for y".

Indeed I can:

$$\text{For all } x \in A,\ g\Big(f(x)\Big) \overset{5.1.22}{=} (gf)(x) \overset{by\ (1)}{=} \mathbf{1}_A(x) = x \tag{2}$$

So to solve, take $y = f(x)$.

Mar. 11, 2024

About f:

Totalness: Start from "$x = g(f(x))$, for each $x \in A$, is **true**" by (2).

This is *the same as* "$x\, f \circ g\, x$ is true" —for all $x \in A$. Therefore, for each such x, there must be a z such that $x\, f\, z$ (and $z\, g\, x$).

Thus f is total on A.

1-1 ness: For the 1-1ness, we prove $f(a) = f(b) = c$ implies $a = b$.

Assume then $f(a) = f(b) = c$ and *apply* g to both sides of **the first** "=", meaning call g with input c.

Under any name the call to c returns the same object. We get $g(f(a)) = g(f(b))$, that is,

$$(gf)(a) = (gf)(b)$$

But this says $a = b$, by $gf = \mathbf{1}_A$, and we are done. □

5.1.27 Example. (READ ME!) *The above is as much as can be proved.* For example, say $A = \{1, 2\}$ and $B = \{3, 4, 5, 6\}$.

Let $f : A \to B$ be $\{(1, 4), (2, 3)\}$ and

$g : B \to A$ be $\{(4, 1), (3, 2), (6, 1)\}$, or in friendlier notation

$f(1) = 4$
$f(2) = 3$
and
$g(3) = 2$
$g(4) = 1$
$g(5) \uparrow$
$g(6) = 1$

Clearly, $gf = \mathbf{1}_A$ holds, but note:
(1) f is not onto B.
(2) g is neither 1-1 nor total.

□ 

5.1.28 Example. (READ ME!) With $A = \{1, 2\}$, $B = \{3, 4, 5, 6\}$ and $f : A \to B$ and $g : B \to A$ as in the previous example, consider also the functions $\tilde{f}$ and $\tilde{g}$ given by

$\tilde{f}(1) = 6$
$\tilde{f}(2) = 3$
and
$\tilde{g}(3) = 2$
$\tilde{g}(4) = 1$
$\tilde{g}(5) = 2$
$\tilde{g}(6) = 1$

Clearly, $\tilde{g}f = \mathbf{1}_A$ and $g\tilde{f} = \mathbf{1}_A$ hold, but note:

(1) $f \neq \tilde{f}$.

(2) $g \neq \tilde{g}$.

Thus, neither left nor right inverses need to be unique. The article "a" in the definition of said inverses was well-chosen. □

The following two *partial converses* of 5.1.25 are useful.

5.1.29 Theorem. *Let $f : A \to B$ be total and 1-1. Then there is an* onto function $g : B \to A$ *such that* $gf = \mathbf{1}_A$.

Proof. Consider the *converse* relation (4.4.14) of f —that is, the *relation* f^{-1}— but call it g instead. I show that this "g" works. So:

$$x\, g\, y \overset{\text{Def}}{\text{ iff }} y\, f\, x \text{ (Says } x\, f^{-1}\, y \text{ iff } y\, f\, x) \tag{1}$$

$$\text{ran}(g) = \text{ran}(f^{-1}) \overset{4.4.15}{=} \text{dom}(f) \overset{f \text{ is total}}{=} A, \text{ so } g \text{ is onto } A \tag{2}$$

By Exercise 5.1.18 (do this!), $g : B \to A$ is a (possibly nontotal) function.

Since f is total on its left field A, we have

$$\textbf{true} : \text{For any } y \in A, \text{ a } x \in B \text{ exists such that } yfx \quad (f(y) = x). \tag{3}$$

By (1) we also have xgy, thus $yfx \wedge xgy$ is true, from which we get

$$\textbf{true} : yf \circ gy, \text{ for any } y \in A \tag{4}$$

Since g is a function, we can write (4) as

$$\textbf{true} : y(gf)y, \text{ for any } y \in A$$

that is,

$$\textbf{true} : (gf)(y) = y, \text{ for any } y \in A \tag{5}$$

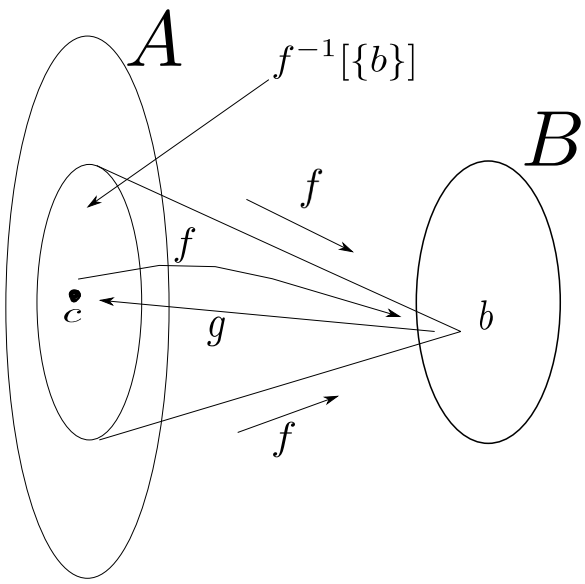
Since (5) is true, for all $y \in A$, it says $gf = \mathbf{1}_A$.

By 5.1.25, (5) proves that g is onto —but we already got this in (2).

We got both statements (one is that g —a left inverse— exists) that we needed to prove. □

5.1.30 Theorem. *Let* $f : A \to B$ *be onto. Then there is a* total and 1-1 $g : B \to A$ *such that* $fg = \mathbf{1}_B$.

Proof. By assumption (ontoness), $\emptyset \neq f^{-1}[\{b\}] \subseteq A$, for all $b \in B$.



To define $g(b)$ **choose** *ONE* —we want g to be single-valued!

$$c \in f^{-1}[\{b\}] \qquad (\dagger)$$

and Define $g(b) = c$.

▶ Do so for *each* $b \in B$. ◀

Since $f(c) = b$ by $(\dagger)$, we get $f(g(b)) = b$ for all $b \in B$, that is, $fg = \mathbf{1}_B$.

Hence g is 1-1 and total by 5.1.25. □

5.1.31 Remark. (Axiom of Choice) The proof of 5.1.30 states

$$\underline{\textbf{choose}}\ \textit{one}\ c \in f^{-1}[\{b\}]$$

and that must be done for all $b \in B$ that may be *infinitely many*.

Choosing once is OK: "**We know $f^{-1}[\{b\}] \neq \emptyset$. So, let $c \in f^{-1}[\{b\}] \neq \emptyset$**".

We can fit inside a proof any finite number of copies of the statement in quotes for various b.

But how do you choose "the" c **for infinitely many b**? If we were dealing with natural numbers I can see that (How?).

But not with the reals and not with arbitrary unspecified sets!

How do you DESCRIBE in a finite mathematical way the process of choosing ONE element out of each of (potentially) infinitely many nonempty sets?

Why finite? Because a proof MUST be written in a finite space of symbols and words!

How —for example (due to Russell)— do you describe the process of choosing ONE sock from each of infinitely many pairs?

True, you might sit there for an infinite amount of time, and pick ONE sock at random from each pair. But can you sit that long? Even if you can, you will end up (when you write all this up using infinite amount of space in your proof. This is NOT allowed!

In set theory one takes as an axiom that a SET of (results of) c-choices exists! They call it the "Axiom of Choice". It says that **if we have an infinite *set* family of nonempty *sets* a set of representatives from each set in the family exists.** □

The Axiom of Choice says that:

if F is a *set* **family** of *nonempty sets*, then a function C exists such for each $A \in F$ we have $C(A) \in A$.

Thus the "mathematical way" to define g in the previous proof — rather than the blabla starting at sign (†)— is simply,

$$g(b) \stackrel{Def}{=} C\Big(f^{-1}[\{b\}]\Big), \text{ for all } b \in B$$

The big red brackets MUST be round! Right?

5.2. *Finite and Infinite Sets*

Broadly speaking (that is, with very little detail contained in what I will say next) we have sets that are *finite* —intuitively meaning that *we can "count"* all *their elements in a "finite*[†] *amount" of "time"* (but see the  -remark 5.2.3 below)— and those that are not, the *infinite* sets!

What is a mathematical way to say all this?

[†]**I know, I know! We cannot define "finite" by assuming I already know what "finite" means. And there is a problem with "time" too!**

Any *counting process* of the elements of a finite set A will have us say out loud —every time we *pick*, or *point* at, an element of A— "0th", "1st", "2nd", etc.,

Once we reach and pick the *last* element of the set, we finally pronounce "nth", for some appropriate n that we reached in our counting (Again, see 5.2.3.)

Thus, mathematically, we *are pairing* each member of the set —or *label* each member of the set— with a member from $\{0, \ldots, n\}$.

Thus the following makes sense:

5.2.1 Definition. (Finite and infinite sets) A set A is *finite* iff it is **either empty, OR** —for some $n \in \mathbb{N}$— is in 1-1 correspondence with $\{x \in \mathbb{N} : x \leq n\}$.

This "normalised" (or "**canonical**") "small" set of natural numbers we usually denote by $\{0, 1, 2, \ldots, n\}$.

If a set is *not* finite, then it is —**by definition**— *infinite*. □

5.2.2 Example. For any n, $\{0, \ldots, n\}$ is finite since, trivially,

$$\{0, \ldots, n\} \sim \{0, \ldots, n\}$$

using the identity (Δ) function on the set $\{0, \ldots, n\}$. □

Mar. 13, 2024

5.2.3 Remark. One must be careful when one attempts to explain finiteness via counting by a human.

For example, Achilles[†] could count *infinitely many objects* by constantly accelerating his counting process as follows:

He procrastinated for a *full second*, and then counted the first element. Then, he counted the second object *exactly after* 1/2 a second from the first. Then he got to the third element $1/2^2$ seconds after the previous, ..., he counted the n th item at exactly $1/2^{n-1}$ seconds after the previous, and so on *forever*.

Hmm! It was *not* "forever", was it? After a total of 2 seconds he was done!

You see (as you can easily verify from your calculus knowledge (limits)),[‡]

$$1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^{n-1}}+\ldots=\frac{1}{1-1/2}=2 \text{ seconds}$$

So "clock-time" is *not* a good determinant of finiteness! □

[†]OK, he was a demigod; but only "demi".

[‡]$1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^{n-1}}=\frac{1-1/2^n}{1-1/2}$. Now let n go to infinity at the limit.

5.2.4 Theorem. (This is Key!) *If* $X \subsetneqq \{0, \dots, n\}$, *then there is NO* onto *function* $f : X \to \{0, \dots, n\}$.


I am saying, *NO such* f, *whether total or not exists*; actually, totalness is *immaterial*.


Proof. First off, the claim *is true* if $X = \emptyset$, since then any such f equals $\emptyset$ —no inputs, therefore no outputs!

The range of f is empty so f cannot be onto any nonempty set.


But how about the case of $X \neq \emptyset$?


Let us proceed by way of contradiction, and assume that the theorem is *wrong*.

That is, **assume that** it *IS* possible to have such onto functions, for *some* n and *well-chosen* $\emptyset \neq X \subsetneqq \{0, \dots, n\}$.

So let n_0 be the *smallest* n that *contradicts* the theorem, and let X_0 be a *corresponding* set "X" that supports the contradiction, that is,

$$X_0 \subsetneqq \{0, \ldots, n_0\} \text{ AND } f : X_0 \to \{0, \ldots, n_0\} \textbf{ IS } \textit{onto} \quad (1)$$

Firstly, we saw that $X_0 \neq \emptyset$, since $X_0 = \emptyset$ *does NOT FAIL the theorem*.

Secondly, $n_0 > 0$, since otherwise —i.e., *IF* $n_0 = 0$— then $X_0 = \emptyset$ (Why?) and, as remarked, the latter *does NOT FAIL the theorem*.

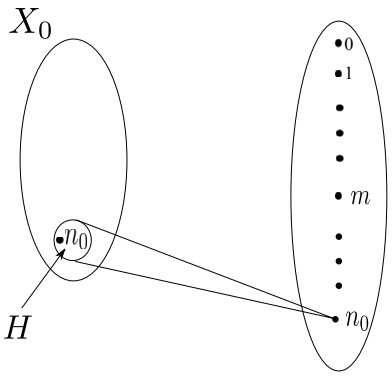
Let us set $H = f^{-1}[\{n_0\}]$, **that is, all inputs that cause output** n_0.

$\emptyset \neq H \subseteq X_0$; the $\neq$ by ontoness.

Case 1. $n_0 \in H$. Then removing all pairs (a, n_0) from f —all these have $a \in H$— we get a new function $f' : X_0 - H \to \{0, 1, \ldots, n_0 - 1\}$, which *is still onto* as we **only removed inputs that cause output** n_0 —and thus contradicts the theorem.

This also contradicts minimality of n_0 since $n_0 - 1$ works too! (works to provide an onto map and thus refute the theorem).

$$H = f^{-1}[\{n_0\}]$$



Case 2. $n_0 \notin H$.

(a) Subcase where $n_0 \notin X_0$ too. Given that $X_0 \subsetneqq \{0, 1, \ldots, n_0\}$, then also $X_0 \subseteq \{0, 1, \ldots, n_0 - 1\}$.[†] By $H \neq \emptyset$, $X_0 - H \subsetneqq \{0, 1, \ldots, n_0 - 1\}$. As in Case 1, $f' : X_0 - H \to \{0, 1, \ldots, n_0 - 1\}$ is onto. Contradiction to minimality of n_0.

NOTE that $f(n_0) \uparrow$ in this case since f has X_0 as left field and $n_0 \notin X_0 \supseteq \text{dom}(f)$.

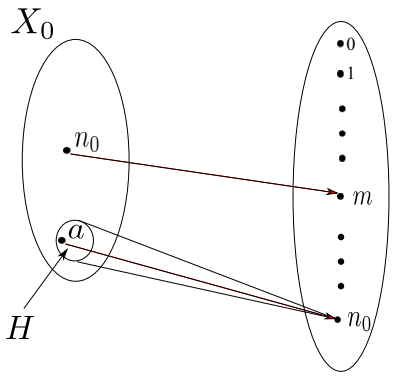
[†] $X_0 \subsetneqq \{0, 1, \ldots, n_0\}$ might be due to $X_0 = \{0, 1, \ldots, n_0 - 1\}$ and $n_0 \notin X_0$.

(b) $n_0 \in X_0$. We have two subcases:

- $f(n_0)\uparrow$. Then we (almost) act as in Case 2(a):
  The new "X_0" is $(X_0 - H) - \{n_0\}$.
  We remove $n_0 \in X_0$ to ensure that the new "X_0" *will* be a subset of $\{0, 1, \ldots, n_0 - 1\}$ and *we get a contradiction exactly per Case 2(a)*. The new onto function is

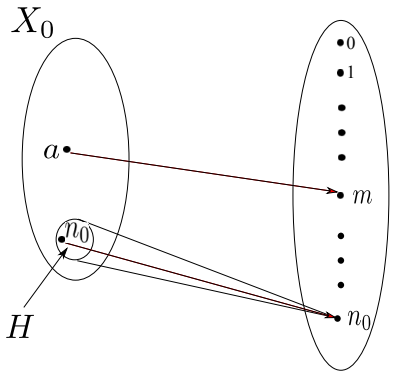
$$f' = f - \underbrace{H \times \{n_0\}}_{remove\ all\ (a,n_0)}$$

- We have the picture below —that is, $f(n_0) = m \neq n_0$ for some m.



We simply transform the picture to the one below, "correcting" f to have $f(a) = m$ and $f(n_0) = n_0$, that is defining a new "f" that we will call f' by

$$f' = \Big(f - \{(n_0, m), (a, n_0)\}\Big) \cup \{(n_0, n_0), (a, m)\}$$



We are back to Case 1 with the function f'. □

5.2.5 Corollary. (Pigeon-Hole Principle) *If* $m < n$, *then* $\{0, \ldots, m\} \not\sim \{0, \ldots, n\}$.

Proof. If the conclusion fails then we have an onto $f : \{0, \ldots, m\} \to \{0, \ldots, n\}$, contradicting 5.2.4. □

Important!

5.2.6 Theorem. *If* A *is finite due to* $A \sim \{0, 1, 2, \ldots n\}$ *then there is* ***no justification of finiteness via another canonical set*** $\{0, 1, 2, \ldots m\}$ *with* $n \neq m$.

Proof. If $\{0, 1, 2, \ldots n\} \sim A \sim \{0, 1, 2, \ldots m\}$, then $\{0, 1, 2, \ldots n\} \sim \{0, 1, 2, \ldots m\}$ by 5.1.20, hence $n = m$, otherwise we contradict 5.2.5. □

5.2.7 Definition. Let $A \sim \{0, \ldots, n\}$. Since n is uniquely determined by A we say that A has $n+1$ elements and write $|A| = n+1$. □

5.2.8 Corollary. *There is no onto function from* $\{0,\ldots,n\}$ *to* $\mathbb{N}$.

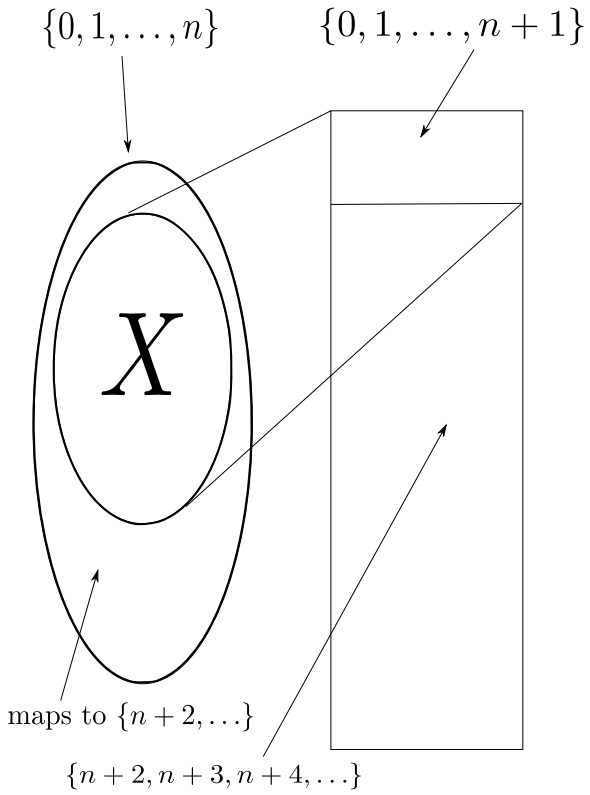
"For all $n \in \mathbb{N}$, there is no…" is, of course, implied.

Proof. Fix an n. By way of contradiction, let $g : \{0,\ldots,n\} \to \mathbb{N}$ be onto.

Let X be the set of all inputs that g maps *onto* $\{0,\ldots,n+1\}$. (†)

$$X \overset{Def}{=} g^{-1}[\{0,1,\ldots,n+1\}] \subseteq \overbrace{\{0,1,\ldots,n\}}^{\text{left field of } g} \subsetneqq \{0,1,\ldots,n,n+1\} \quad (‡)$$

As (‡) entails $X \subsetneqq \{0,\ldots,n+1\}$, using (†) we have contradicted Theorem 5.2.4



□

5.2.9 Corollary. $\mathbb{N}$ *is infinite.*

Proof. By 5.2.1 the opposite case requires that there is an n and a function $f : \{0, 1, 2, \ldots, n\} \to \mathbb{N}$ that is a 1-1 correspondence. *Impossible*, since any such an f will fail to be *onto* $\mathbb{N}$. □

Our mathematical definitions have led to what we hoped they would:

For example, that $\mathbb{N}$ *is* infinite as we intuitively understand, notwithstanding Achilles's accelerated counting!

$\mathbb{N}$ is a "canonical" infinite set that we can use to *index* or *label* the members of many infinite sets.

Sets that can be indexed using natural number indices

$$a_0, a_1, \ldots$$

are called *countable*.

Wait! I said "sets". Is that legitimate?

In the interest of *technical flexibility*, *we do not insist* that *all* members of $\mathbb{N}$ be used as indices.

We might enumerate with *gaps*:

$$b_5, b_9, b_{13}, b_{42}, \ldots$$

Thus, informally, a set A is *countable* if it is empty or (in the opposite case) if there is a way to index, hence enumerate, all its members in an array, utilising indices from —but **not necessarily all**— $\mathbb{N}$. *See also 5.1.8 regarding indexing/labelling.*

It *is* allowed to *repeatedly list any element of A*, so that finite sets *are* countable.

Mar. 15, 2024

For example, the set $\{42\}$:

$$42, 42, 42, \overbrace{\ldots}^{42 \text{ forever}}$$

We may think that the enumeration above is done by assigning to "42" *all* of the members of $\mathbb{N}$ as indices, in other words, the enumeration is effected, for example, by the constant function $f : \mathbb{N} \to \{42\}$ given by $f(n) = 42$ for all $n \in \mathbb{N}$.

This is consistent with our earlier definition of indexing (5.1.8).

Now, mathematically,

5.2.10 Definition. (Countable Sets) We call a set A *countable* if there is an *onto* function $f : \mathbb{N} \to A$.

We *do NOT* require f to be *total*.

But, $\emptyset$, the empty function, is onto $\emptyset$, the empty set.

Thus the definition makes $\emptyset$ countable.

If $f(n)\downarrow$, then we say that $f(n)$ is the nth element of A in the enumeration f.

We often write f_n instead of $f(n)$ and then call n a "subscript" or "index". □

Thus a set is countable iff it is the *range* of some function that has $\mathbb{N}$ as its *left field*.

Some set theorists also define sets that can be enumerated using *all* the elements of $\mathbb{N}$ as indices *without repetitions*.

5.2.11 Definition. (Enumerable or denumerable sets) A set A is *enumerable* iff $A \sim \mathbb{N}$ iff $\mathbb{N} \sim \mathbb{A}$. □

5.2.12 Example. Every enumerable set is countable, but the converse fails. For example, $\{1\}$ is countable but not enumerable due to 5.2.8.

$\{2n : n \in \mathbb{N}\}$ is enumerable, with $f(n) = 2n$ effecting the 1-1 correspondence $f : \mathbb{N} \to \{2n : n \in \mathbb{N}\}$. □

5.2.13 Theorem. *If A is an infinite subset of $\mathbb{N}$, then $A \sim \mathbb{N}$. That is, A is enumerable.*

Proof. We will build a 1-1 and total enumeration of A, presented in a finite manner as a (pseudo) program below, which enumerates all the members of A in strict ascending order and arranges them in an array

$$a(0), a(1), a(2), \ldots a(k-1), \ldots \tag{1}$$

n	$\leftarrow 0$	
$a(0)$	$\leftarrow \min A$	Initialisation; $A \neq \emptyset$
while	$A - \{a(k) : k \leq n\} \neq \emptyset$	
$a(n+1)$	$\leftarrow \min \Big(A - \{a(k) : k \leq n\}\Big)$	
n	$\leftarrow n+1$	
end while		

Note that the sequence $\{a(0), a(1), \ldots, a(m)\}$ is strictly increasing for any m. Indeed (instruction below the word "**while**"),

$$a(n+1) = \min \Big(A - \{a(0), a(1), \ldots, a(n)\}\Big)$$

hence,

$$a(0) < a(1),\ a(0) < a(1) < a(2), \ldots, \overbrace{\underbrace{a(0) < a(1) < \cdots < a(n)}_{\text{all these, selected earlier, are } < a(n+1)}}^{\text{say we verified ordering up to } a(n)} < a(n+1)$$

Will this loop ever exit?

Suppose yes. Then, say, this happens the first time we got $A - \{a(k) : k \leq n\} = \emptyset$ for some n, that is, $A = \{a(0), a(1), \ldots, a(n)\}$.

The function a taking $\{0, 1, \ldots, n\}$ onto A (why "onto"?) is total on $\{0, 1, \ldots, n\}$ and strictly increasing, so is 1-1. Thus $A \sim \{0, 1, \ldots, n\}$ and A is finite. A contradiction.

Thus, we never exit the loop! We do obtain for each n an entry to put in "$a(n)$"

This says that the function $n \mapsto a(n)$ is defined for every n: In other words, it is total!

Now, distinct inputs cause distinct outputs in the function $n \mapsto a(n)$ since the function satisfies $a(i) < a(i+1)$ for all i.

Thus the function is 1-1.

The function $n \mapsto a(n)$ is also *onto* A, so all in all we got $\mathbb{N} \sim A$ via a.

Wait! Why is $n \mapsto a(n)$ onto?

If you don't think so, let $m \in A$ be one entry we missed *and did not insert in the array* a.

Let n be *the smallest* such

$$m < a(n) \qquad (\dagger)$$

Such an n exists since

$$\ldots, a(i) < a(i+1), \ldots$$

is a strictly increasing sequence of natural numbers that goes on forever —the entries $a(i)$ get larger and larger (by at least a step of plus-1 from the previous entry) with no end.

At the step at which I select $a(n)$ both it —I did not select it yet— and m —I never selected it— are in the residual A.

But we selected $a(n)$ at this step and yet m is smaller. Contradiction!

So no "forgotten" m (as in $(\dagger)$) exists. The set of entries of the array a does equal A, or, $n \mapsto a(n)$ is onto A. □

5.2.14 Theorem. *Every infinite countable set A is enumerable.*

Proof. Let $f : \mathbb{N} \to A$ be onto, where A is infinite.

 Reminder: f need not be total. 

Let $g : A \to \mathbb{N}$ such that $fg = \mathbf{1}_A$ (5.1.30).

Thus, g is *total* and 1-1 and moreover is *onto* $B = \text{ran}(g)$.

 **Because**: *Every function* is *onto its range!* 

We have the following configuration:

$$A \overset{g}{\underset{\sim}{\rightarrow}} B \subseteq \mathbb{N} \xrightarrow{f} A \tag{1}$$

1. Hence B is *infinite*, else we would have $A \sim B \sim \{0, \ldots, n\}$ (some n) and thus $A \sim \{0, \ldots, n\}$ (see Exercise 5.1.20) making A *finite*!

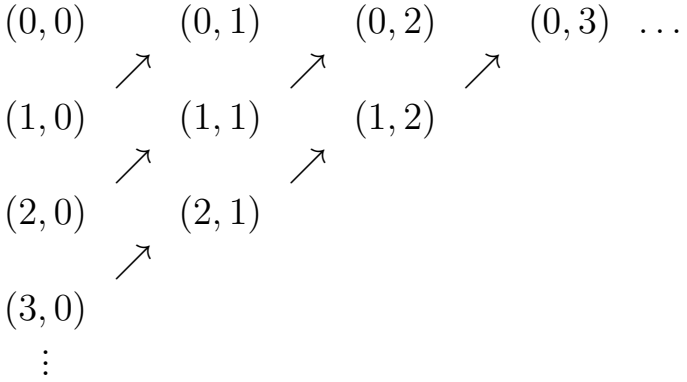
2. It follows by 5.2.13 that $B \sim \mathbb{N}$ hence $A \sim \mathbb{N}$ via $A \sim B \sim \mathbb{N}$ and 5.1.20 once more. □

So, if we can enumerate an infinite set at all, then we can enumerate it without repetitions.

Mar. 18, 2024

5.2.15 Example. (Important) We can linearise an infinite square matrix of elements in each location (i, j) by devising a traversal that will go through each (i, j) entry *once*, and will *not miss any entry*!

In the literature one often sees the method diagrammatically, see below, where arrows *clearly* indicate the sequence of traversing, with the understanding that we use the arrows by picking the first unused chain of arrows from left to right.



So the linearisation induces a 1-1 correspondence between $\mathbb{N}$ *and the linearised sequence of matrix entries, that is, it shows that* $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$.

□

In short,

5.2.16 Theorem. *The set* $\mathbb{N} \times \mathbb{N}$ *is countable. In fact, it is enumerable.*

Is there a "mathematical" way to do this? Well, the above IS mathematical, don't get me wrong, but is given in *outline*. It is kind of like an argument in geometry, where we rely on drawings (figures).

READ ME! Here are the "algebraic" details:

Proof. (of 5.2.16 with an "algebraic" argument). Let us call $i+j+1$ the "*weight*" of a pair (i,j). The weight is the number of elements in the group:

$$(i+j,0),(i+j-1,1),(i+j-2,2),\ldots,(i,j),\ldots,(0,i+j)$$

Thus the diagrammatic enumeration proceeds by enumerating *groups* by increasing weight

$$1,2,3,4,5,\ldots$$

and in each group of weight k we enumerate in *ascending order of the second component.*

Thus the (i,j) th entry occupies position j *in its group* —the first position in the group being the 0 th, e.g., in the group of $(3,0)$ the first position is the 0 th— and this position *globally* is the number of elements in all groups *before* group $i+j+1$, *plus* j. Thus the first available position for the first entry —$(i+j,0)$— of group (i,j) members is just after this many occupied positions:

$$1+2+3+\ldots(i+j)=\frac{(i+j)(i+j+1)}{2}$$

That is,

$$\text{global position of } (i,j) \text{ is this: } \frac{(i+j)(i+j+1)}{2}+j$$

The function f which for all i,j is given by

$$f(i,j)=\frac{(i+j)(i+j+1)}{2}+j$$

is the algebraic form of the above enumeration. □

There is an easier way to show that $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$ without diagrams:

By the unique factorisation of numbers into products of primes (Euclid) the function

$g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ given for all m, n by $g(m, n) = 2^m 3^n$ is 1-1, since Euclid proved that $2^m 3^n = 2^{m'} 3^{n'}$ implies $m = m'$ and $n = n'$.

It is not onto as it never outputs, say, 5, but $\text{ran}(g)$ is an *infinite* subset of $\mathbb{N}$ (Exercise!).

Thus, trivially,

$$\mathbb{N} \times \mathbb{N} \overset{via\ g}{\sim} \text{ran}(g) \sim \mathbb{N}$$

the 2nd "$\sim$" by 5.2.13. END READ ME!

5.2.17 Exercise. Say $A \subseteq B$ and A is infinite. Prove that B too is infinite. □

5.2.18 Exercise. If A and B are enumerable, so is $A \times B$.

Hint. So, $\mathbb{N} \sim A$ and $\mathbb{N} \sim B$. Can you show now that $\mathbb{N}\times\mathbb{N} \sim A\times B$?

□

With little additional effort one can generalise to the case of $\bigtimes_{i=1}^{n} A_i$.

5.2.19 Remark.

1. Let us collect a few more remarks on countable sets here. Suppose now that we start with a countable set A. Is every subset of A countable?

   Yes, because the composition of onto functions is onto. **Exercise!**

   5.2.20 Exercise. *What does composition of onto functions have to do with this?* Well, prove that if $B \subseteq A$ then there is a *natural* onto function $g : A \to B$. Which one? Now study the *Hint*.

   Hint. Think "natural"! Get a *natural* total and 1-1 function $h : B \to A$ to obtain (via 5.1.30) an onto $g : A \to B$. Then use the onto $f : \mathbb{N} \to A$ (A is countable) to get the onto $gf : \mathbb{N} \to B$ to settle Exercise 1. above. □

2. As a special case, **if A is countable, then so is $A \cap B$ for any B**, since $A \cap B \subseteq A$.

3. How about $A \cup B$? If both A and B are countable, then so is $A \cup B$. Indeed, and without inventing a new technique, let

$$a_0, a_1, \ldots$$

be an enumeration of A and

$$b_0, b_1, \ldots$$

for B. Now form an infinite matrix with the A-enumeration as the 1st row, while each remaining row is the same as the B-enumeration. Now linearise this matrix!

Of course, we may alternatively adapt the unfolding technique to an infinite matrix of just two rows. **How?**

... OR, just use the "common sense" enumeration back and forth between the "a_i's" and the "b_i's":

$$a_0, b_0, a_1, b_1, a_2, b_2, a_3, b_3, \ldots$$

4. **5.2.21 Exercise.** Let A be enumerable and an enumeration of A

$$a_0, a_1, a_2, \ldots \tag{1}$$

is given.

So, this is an enumeration without repetitions.

Use techniques we employed in this section to propose a new enumeration in which *every* a_i is listed *infinitely many times* (this is useful in some applications of logic). □

5.2.22 Example. Any proper subset X of $\{0, 1, \ldots, n\}$ —any $n \geq 0$— is finite.

Say X is *infinite* instead. Since $X \subseteq \{0, 1, \ldots, n\} \subseteq \mathbb{N}$, we have (5.2.13) $X \sim \mathbb{N}$, that is, X is *enumerable*.

So let $f : X \to \mathbb{N}$ be 1-1, total (on X) and onto. Then (the possibly nontotal) $f' : \{0, \ldots, n\} \to \mathbb{N}$ given, for all $x \in \{0, 1, \ldots, n\}$, by

$$f'(x) = \begin{cases} f(x) & if\ x \in X \\ \uparrow & if\ x \notin X \end{cases}$$

is onto, contradicting 5.2.8. □

5.3. Diagonalisation and uncountable sets

5.3.1 Example. Suppose we have a 3×3 matrix

$$\begin{matrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{matrix}$$

and we are asked:

Find a sequence of three numbers, *using only* 0 *or* 1, that does not *fit* as a row of the above matrix —i.e., is *different from all rows*.

Sure, you reply: Take 1 1 1. Or, take 0 0 0.

Both are correct.

But what if the matrix were big, say, $10^{350000} \times 10^{350000}$, or even *infinite*?

Is there a *finitely describable technique* that can produce an "unfit" row for *any* square matrix, even an *infinite* one?

□

Yes, it is Cantor's *diagonal method* or technique.

5.3.2 Definition. (Diagonalisation: How-to) Cantor noticed that *any row that fits in a square matrix* M as the, say, i-th row, *intersects* the *main diagonal* at entry $M(i,i)$.

Why?

$$\text{Row } i : M(i,0), M(i,1), M(i,2), \ldots, \overbrace{M(i,i)}^{i\text{-th member of row}}, M(i,i+1), \ldots$$

Thus if we take the main diagonal —*a sequence that has the same length as any row*— and *make a copy of it changing every one of the original entries* $M(x,x)$ to a different one

$$\overline{M(x,x)}$$

then this changed copy (of the main diagonal) will *not* fit *anywhere in* M as a row!

Note that the *Main (Original) Diagonal* is the *sequence* of entries below:

$$\begin{array}{cccccc} pos.\ 0 & pos.\ 1 & pos.\ 2 & & pos.\ i & \\ \downarrow & \downarrow & \downarrow & & \downarrow & \\ M(0,0), & M(1,1), & M(2,2), & \ldots, & M(i,i), & \ldots \end{array}$$

The modified diagonal is (where we named "D" the array below):

$$\begin{array}{cccccc} & pos.\ 0 & pos.\ 1 & pos.\ 2 & & pos.\ i \\ & \downarrow & \downarrow & \downarrow & & \downarrow \\ D = & \overline{M(0,0)}, & \overline{M(1,1)}, & \overline{M(2,2)}, & \ldots, & \overline{M(i,i)}, \ldots \end{array}$$

where, *for all positions* i, $\overline{M(i,i)} \neq M(i,i)$.

Thus if D *were to fit as row* x*, then the* x*-th element of* D —$\overline{M(x,x)}$— *would overlap the (original)* x*-th element of the matrix* M —$M(x,x)$.

But these two are *different*!

So, the modified diagonal D does NOT FIT as the x-th row! □

This HOW TO would give the alternative answer 0 1 0 to our original question in 5.3.1.

5.3.3 Example. We have an infinite *matrix* M of 0-1 entries. Can we construct a row-long *array* of 0-1 entries that does not match *any* row in the matrix?

Yes, to get the counterpart of D above just define for all x:

$$\overline{M(x,x)} = 1 - M(x,x)$$

In words, take the main diagonal and *flip every entry* (0 to 1; 1 to 0).

Now refer to 5.3.2. □

5.3.4 Example. (Cantor) Let S denote the set of **all** *infinite sequences* —also called *infinite* **strings**— of 0s and 1s.

Pause. What is an *infinite sequence*?
It is a total function f on $\mathbb{N}$ (left field) that we want to view as **the array of its outputs**:

$$f(0), f(1), f(2), \ldots, f(n), \ldots \tag{1}$$

(1) is an infinite sequence of 0s and 1s if $\text{ran}(f) = \{0, 1\}$.

We say that "the n-th member of the sequence is $f(n)$".◀

Can we arrange *ALL* of S in an *infinite matrix* —one element per row?

No, since the preceding example shows that we would miss at least one infinite sequence ROW (i.e., we *would fail to list it as a row*), because a sequence of infinitely many 0s and/or 1s can be found, that does not match ANY row!

□

5.3.5 Definition. (Uncountable Sets) A set that is *not* countable is called *uncountable*. □

If it is *not* countable —is *uncountable*— then it is *NOT* enumerable, **right?**

Example 5.3.4 shows that uncountable sets exist. Here is a more interesting one.

5.3.6 Example. (Cantor) The set of real numbers in the interval

$$(0,1) \stackrel{\text{Def}}{=} \{x \in \mathbb{R} : 0 < x < 1\}$$

is uncountable. This is done via an elaboration of the argument in 5.3.4.

Think of a member of $(0,1)$, *in form*, as an infinite sequence of numbers from the set $\{0,1,2,3,4,5,6,7,8,9\}$ prefixed with a dot; that is, think of the number's decimal notation.

Some numbers have representations that end in 0s after a certain point. We call these representations *finite*. Every such number has also an "*infinite representation*" since the non zero digit d immediately to the left of the infinite tail of 0s can be converted to $d-1$ followed by an infinite tail of 9s, without changing the value of the number.

We allow only infinite representations.

Assume now **by way of contradiction** that a listing of all members of $(0,1)$ exists, *listing them via their infinite representations* —where the leading decimal point is omitted and all a_{ij} satisfy $0 \leq a_{ij} \leq 9$ (decimal digits).

$$\begin{array}{l} a_{00}a_{01}a_{02}a_{03}a_{04}\ldots \\ a_{10}a_{11}a_{12}a_{13}a_{14}\ldots \\ a_{20}a_{21}a_{22}a_{23}a_{24}\ldots \\ a_{30}a_{31}a_{32}a_{33}a_{34}\ldots \\ \vdots \end{array} \tag{1}$$

The "How To" of Definition 5.3.2 is applied now to obtain a

number

$$D = (.)\overline{a_{00}}\,\overline{a_{11}}\,\overline{a_{22}}\ldots\overline{a_{xx}}\ldots$$

where

$$\overline{a_{xx}} = \begin{cases} 2 & \text{if } a_{xx} = 0 \vee a_{xx} = 1 \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

Clearly (by 5.3.2) D does not fit in *any row i of (1)*, that is, the number it represents *is both*

- *In* $(0,1)$ —since its digits are 1 or 2 it is $0 < D < 1$,

  AND

- *Not in* $(0,1)$ —by the diagonalisation in (2).

This contradiction shows that we do *NOT* have the *enumeration* of all of $(0,1)$ depicted as (1): *The real interval is uncountable*. □

5.3.7 Example. (5.3.4 Revisited) Consider the set of *all* total functions from $\mathbb{N}$ to $\{0,1\}$. Is this countable?

Connection with 5.3.4? Well, a total function f with right field $\{0,1\}$ is an infinite 0-1 string

$$f = f(0), f(1), f(2), \ldots, f(i), \ldots$$

So, to fit all such strings in a matrix —which 5.3.4 says is impossible— is the same as asking whether we can fit all total functions f with $\{0,1\}$ as right field in an enumeration $f_0, f_1, \ldots$.

If so, each f_i is a "header" of a row of said matrix:

f_0=$f_0(0)f_0(1)f_0(2)f_0(3)\ldots$
f_1=$f_1(0)f_1(1)f_1(2)f_1(3)\ldots$
$\vdots$
f_i=$f_i(0)\, f_i(1)\, f_i(2)\, f_i(3)\ldots$
$\vdots$

Here is a *direct proof* of the uncountability of all total f with $\{0,1\}$ as right field:

If there *IS* an enumeration of these one-variable functions

$$f_0, f_1, f_2, f_3, \ldots \tag{1}$$

consider the function $g : \mathbb{N} \to \{0,1\}$ given by $g(x) = 1 - f_x(x)$.

Clearly, this *must* appear in the listing (1) since it has the correct left and right fields, and is total.

Too bad! If $g = f_i$ then $g(i) = f_i(i)$. By definition, also $g(i) = 1 - f_i(i)$.

So, $f_i(i) = 1 - f_i(i)$ which is false for total f_i.

A contradiction. □

The *same* argument as above shows that the set of all functions from $\mathbb{N}$ to $\mathbb{N}$ is uncountable.

Taking $g(x) = f_x(x) + 1$ also works here to "systematically change the diagonal" $f_0(0), f_1(1), \ldots$ since we are not constrained to keep the function values in $\{0, 1\}$.

Mar. 20, 2024

5.3.8 Example. (Cantor) What about the set of all subsets of $\mathbb{N}$ — $\mathcal{P}(\mathbb{N})$ or $2^{\mathbb{N}}$?

Cantor showed that this is uncountable as well: If not, we have an enumeration of all its members as

$$S_0, S_1, S_2, \ldots \tag{1}$$

Define the set

$$D \overset{Def}{=} \{x \in \mathbb{N} : x \notin S_x\} \tag{2}$$

So, $D \subseteq \mathbb{N}$, thus it must appear in the list (1) as an S_i: $D = S_i$.

But then

$$i \in D \text{ iff } i \in S_i$$

by virtue of $D = S_i$.

However, also $i \in D$ iff $i \notin S_i$ by Definition (2).

So,

$$i \in S_i \text{ iff } i \in D \text{ iff } i \notin S_i$$

This contradiction establishes that a *legitimate subset of* $\mathbb{N}$, *namely* D, *is* not *an* S_i.

That is, $2^{\mathbb{N}}$ *cannot* be so enumerated; it is uncountable. □

Chapter 6

A Short Course on Predicate (also called "*First Order*") Logic

We have become comfortable in using informal logic in our arguments about aspects of discrete mathematics, in particular proving statements like $\mathbb{A} \subseteq \mathbb{B}$ and $\mathbb{X} = \mathbb{Y}$, for any classes that we know something about their properties/definitions.

Although we have used quantifiers already —$\exists$ and $\forall$— we did so mostly viewing them as *symbolic abbreviations* of *English texts* about mathematics.

In this chapter we will expand our techniques in logic, extending them to include **the correct syntactic —also called "formal"— manipulation of quantifiers.**

This chapter also includes a section on the WHAT and the HOW TO of the versatile *Induction* —or *mathematical induction*— technique used to prove properties of the natural numbers.

We know how to detect fallacious statements formulated in Boolean logic: Simply show by a truth table that the statement is not a tautology (or not a so-called *tautological implication*).

Correspondingly, we will show in the domain of quantifier logic not only how to *prove* statements that include quantifiers but also how to *disprove* false statements that happen to include quantifiers.

6.1. Enriching our proofs to manipulate quantifiers

Manipulation of quantifiers boils down to two questions:

> "*how can I remove a quantifier from the beginning of a formula?*"
> and
> "*how can I add a quantifier at the beginning of a formula?*"

Once we learn these two techniques we will be able to reason within mathematics with ease.

But first let us define once and for all what a **mathematical proof** *looks like*: its *correct*, *expected* *syntax* or *form*.

We will need several Preliminaries: In particular, **new syntactic** concepts and notation to begin with.

1. The alphabet and structure of Predicate Logic formulas.

> Formulas are *strings "over" —meaning, using symbols from— said alphabet* that name statements of mathematics and computer science.

The alphabet of *symbols* that we use to write down formulas contain, **at a minimum**,

$=, \neg, \wedge, \vee, \rightarrow, \equiv, (,), \forall, \exists,$[†] object variables:[‡] $x, y, z, u, v, w, x''_{13} \ldots$

 Among object variables we allow any capital letters as well, with or without primes or subscripts: $Q'''_{12300042}$ 

[†] $\exists$ is introduced as an abbreviation of something more complex in 6.3.2.
[‡] That is, variables that denote *objects* such as numbers, arrays, matrices, sets, trees, etc.

2. One normally works in a **mathematical area of interest**, or *mathematical theory* —such as Geometry, Set Theory, Number Theory, Algebra, Calculus, Theory of Computation— where one needs *additional symbols* to write down formulas, like

$$0, \emptyset, \in, \subseteq, \subsetneq, \bigcap, \bigcup, \cup, \int, \circ, +, \times, \mu$$

and many others.

3. **SYNTAX??** Mathematicians as a rule get to recognise **and** use the *formulas (which* **NAME** *statements)* and *terms (which* **NAME** *objects)* in the math areas of their interest **via practise** without being necessarily taught the recursive definition of the **syntax** of these.

   We will not spell out the syntax in these notes either (but see [Tou08] if you want to know!)

- **Terms** are **OBJECTS** such as:
  (a) variables or
  (b) constants or
  (c) "**function calls**", in the jargon of the computer savvy person.
  These calls take math objects as *inputs* and return math objects as *outputs*.

Examples of Terms are: $\overbrace{x, A, \emptyset, 0, \sqrt{2}, 42}^{var\ or\ const}$,

$$\underbrace{x+y, x\times 3, 0\times x+1, A\cap B}_{calls}$$

NOTE. One is told that $\times$ is stronger than $+$, so, notwithstanding the bracket-parsimonious notation "$0\times x+1$", we know it means "$(0\times x)+1$", so this call returns 1, no matter what we plugged into x.

- **Formulas** are **STATEMENTS**.
  These are *also* function calls, but their output is *restricted* to be one or the other of the truth values true or false (**t** or **f**) but nothing else! Their input, just as in the case for terms, is any math object.

  Examples are:

  $2 < 3$ (**t**),

  $(\forall x)x = x$ (**t**),

  $(\forall x)x = 0$ (**f**),

  $(\exists x)x = 0$ (**t**),

  $x = 0$ neither true nor false; the answer depends on the input we place in x.

  More: $x = x$ (**t**) answer is independent of input.

  $x = 0 \to x = 0$ (**t**) answer is independent of input;

  $x = 0 \to (\forall x)x = 0$ neither true nor false; answer depends on the input in (the leftmost) x!

  The input variable above is the *leftmost* x; the other two (x's)are *bound by "$(\forall x)$"* and *unavailable* to accept inputs. See below.

- If an **occurrence** of a formula variable *is* available for input it could rightly be called "an occurrence as an input variable".

However, such occurrences are instead called *FREE occurrences* in the literature.

> *Non-input occurrences* of a variable are called "**bound**".
>
> Let's *emphasise*: It is not a *variable* x that is free or bound in a formula, but it is *the occurrences of said variable* that we are speaking of, as the immediately preceding example makes clear.

4. In $(\forall x)x = 0$ the variable x is non input, it is "*bound*" we say.

   Just like this: $\Sigma_{i=1}^{4} i$, which means $1 + 2 + 3 + 4$ and **"i" is an illusion!** *NOT* available for input:

   Something like $\Sigma_{3=1}^{4} 3$ is nonsense!

   Similar comment for $(\forall x)x = 42$. Neither of these two occurrences of x is free (= available) for substitution in it.

   No wonder "bound" variables are sometimes called "**apparent** variables".

5. We call $\forall, \exists, \overbrace{\neg, \wedge, \vee, \rightarrow, \equiv}^{Boolean}$ the "*logical connectives*".

6. People avoid cluttering notation with too many brackets by agreeing that the *first 3 connectives* have the same "strength" or "priority"; the *highest*. The remaining connectives have priorities decreasing as we walk to the right.

   Thus, if A and B are (*denote*) formulas, then $\neg A \vee B$ means $(\neg A) \vee B$; $\neg$ wins the "fight" (with $\vee$) for A. If we want $(\forall x)$ to apply to the entire $A \to B$ we must write $(\forall x)(A \to B)$.

   What about $A \to B \to C$ and $A \equiv B \equiv C$? Brackets are **implied** *from right to left*: $A \to (B \to C)$ and $A \equiv (B \equiv C)$.

   And this? $(\exists y)(\forall x)\neg A$. Brackets are *implied, again, from right to left*: $(\exists y)\big((\forall x)(\neg A)\big)$.

BTW, the part of a formula where a $(\forall x)$ or $(\exists x)$ *acts upon* — the "$(\ldots)$" in $(\forall x)(\ldots)$ and $(\exists x)(\ldots)$— is called their *scope*.

Mar. 22, 2024

6.2. Boolean Block Structure

A formula of mathematics may have some *Boolean block structure*. This structure **abstracts** —meaning, **removes detail** to make things **easier**— the original formula **in the hope that** the abstracted one (the "**abstraction**") can be dealt with with Boolean Methods using Table 2.1.

6.2.1 Example. $\boxed{x=0} \to \boxed{x=0} \vee \boxed{z>w}$ [†] has **a** (not "the") Boolean abstraction, or "*Boolean shape*",

$$S_1 \to S_1 \vee S_2 \tag{1}$$

which —as we know from Remark 2.3.4— means $S_1 \to (S_1 \vee S_2)$ since $\vee$ is stronger than $\to$ (in priority).

We then easily find by using Table 2.1 on p.42 that —regardless of the assumed truth values of the blocks, that is, the statements S_1, S_2 and S_3— the truth value of $S_1 \to (S_1 \vee S_2)$ is **always true**.

Such formulas that are true regardless of the truth values of the "blocks" in some chosen Boolean block structure are called **tautologies**.

Thus the *special case* of the "shape" (1) above, namely,

$$x=0 \to x=0 \vee z>w$$

IS a tautology of Predicate Logic. □

[†] The boxes $\boxed{\ }$ are **not** part of the formula; they indicate "boxing".

6.2.2 Example. By contrast $x = x$ is NOT a tautology since it has no Boolean structure: **NO Boolean connectives in $x = x$.** All I can do is to think of $x = x$ as "S_1" —a statement— whose truth value I **cannot decide with Boolean methods.**

On the other hand, invoking the philosophically founded belief (accepted in mathematics) that "every object equals itself" we **can** evaluate $x = x$ **IN Predicate Logic** as true, no matter the "value" —i.e., object assigned to— x. □

6.2.3 Example. Boolean abstractions of a first order formula **are not unique**.

Consider $(\forall x)A \to B$. It has a Boolean structure denoted by the boxing $\boxed{(\forall x)A} \to \boxed{B}$.

This particular abstraction has the shape $S_1 \to S_2$. We cannot conclude that it is a tautology since **letting the first box to be t and the second one to be f** we obtain an overall truth value of false (**f**).

We should not be quick to blame the formula $(\forall x)A \to B$ as the culprit who denies us a tautology. We may need to **find a finer, more sophisticated, Boolean abstraction for it**. *Read on*!

Maybe we are lucky and upon further inspection we find that B has the form $x = 0 \to x = 0$. With this fact uncovered, we propose a new, **refined**, block structure

$$\underbrace{\boxed{(\forall x)A}}_{box\ 1} \to \overbrace{\Big(\underbrace{\boxed{x=0}}_{box\ 2} \to \underbrace{\boxed{x=0}}_{box\ 3}\Big)}^{\mathbf{t}}$$

Under this abstraction the formula is **always true** regardless of the assumed truth values of the three boxes. It is a tautology!

Of course, the only Boolean abstraction possible for $(\forall x)A$ is $\boxed{(\forall x)A}$ since this formula **has no Boolean structure**.

Any Boolean connectives in A are hidden under lock and key in the scope of the shown $(\forall x)$.

All we can say is that, "Boolean-wise", it is a statement S_1 that is NOT always verifiable as true **by Boolean Methods**. **Not a tautology**. □

Tautologies of various shapes play an important role in Predicate Logic proofs.

We write $\models_{taut} A$ to say "A is a tautology" symbolically.

6.2.4 Example.

1. $(\forall x)A$ is *not* a tautology as its abstraction has *two* possible truth values (single "box"; no (visible) Boolean connectives).
2. $x = x$ is *not* a tautology (single "box"; no (visible) Boolean connectives).
3. $x = 0 \to x = 0$ *is* a tautology. □
4. $(\forall x)x = 0 \to x = 0$ is *not* a tautology. The ("the"?!) Boolean abstraction is obtained via the block structure $\boxed{(\forall x)x = 0} \to \boxed{x = 0}$ is **NOT** "always true".

   On the other hand, If we DID evaluate $(\forall x)x = 0$ over the integers $\mathbb{Z}$, then we would see that it is false.[†]

   So the implication is true over the integers!

But we **DON'T evaluate within predicate logic** *when we look for a tautology.*

Why? **Because tautologies are a Boolean phenomenon**! We cannot discover tautologies with predicate logic tools.

□

[†]If we are doing our mathematics restricted to the set $\{0\}$, then, in this "theory" the formula **IS** true!

6.2.5 Definition. (Important! Tautological implication)

We say that the formulas $A_1, A_2, \ldots, A_n$ *tautologically imply* a formula B —in symbols $A_1, A_2, \ldots, A_n \models_{taut} B$— meaning

"the truth of $A_1 \wedge A_2 \wedge \ldots \wedge A_n$ implies the truth of B"

that is, by the truth table for $\rightarrow$, that

$A_1 \wedge A_2 \wedge \ldots \wedge A_n \rightarrow B$ is a tautology

□

So, $\models_{taut}$ *propagates* truth from left to right.

NOTE that **if any of the A_i is f, then NO work is needed to prove the validity of the tautological implication!**

Thus, **Practically**, to prove $A_1, \ldots, A_n \models_{taut} B$ we just **assume** that ALL the A_i are true and then **prove** that B is true.

6.2.6 Example. Here are some easy and some involved tautological implications. They can all be verified using truth tables, either building the tables in full, or taking shortcuts.

1. $A \models_{taut} A$
2. $A \models_{taut} A \vee B$
3. $A \models_{taut} B \to A$
4. $A, \neg A \models_{taut} B$ —any B. Because I do "*work*" only if $A \wedge \neg A$ is true! Just look at 6.2.5 and say: **This says that $A \wedge \neg A \to B$ is "always" t since $A \wedge \neg A$ is always f.**
5. $\mathbf{f} \models_{taut} B$ —any B. Because I do *work* only if lhs is true! See 4. above.
6. Is this a valid tautological implication? $B, A \to B \models_{taut} A$, where A and B are distinct.

   No, for if A is false and B is true, then the lhs is true, but the rhs is false!
7. Is this a valid tautological implication? $A, A \to B \models_{taut} B$? Yes! Say $A = \mathbf{t}$ and $(A \to B) = \mathbf{t}$. Then, from the truth table of $\to$, it *must* be $B = \mathbf{t}$.
8. How about this? $A, A \equiv B \models_{taut} B$? Yes! Verify!
9. **READ ME!** How about this? $A \vee B \equiv B \models_{taut} A \to B$? Yes! I verify:

   First off, **assume** lhs of $\models_{taut}$ —that is, that $A \vee B \equiv B$— is true.

   Two cases:

   - $B = \mathbf{f}$. Then I need the lhs of $\equiv$ to be true to satisfy the red "assume". So $A = \mathbf{f}$ as well and clearly the rhs of $\models_{taut}$ is true with these values.

- $B = \mathbf{t}$. Then I need not worry about A on the lhs. The rhs of $\models_{taut}$ is true by truth table of $\rightarrow$.

10. $A \wedge (\mathbf{f} \equiv A) \models_{taut} B$, for any B. Well, just note that the lhs of $\models_{taut}$ is $\mathbf{f}$ so we need to do no work with B to conclude that the implication is valid.

11. $$A \rightarrow B, C \rightarrow B \models_{taut} A \vee C \rightarrow B$$
This is nicknamed "proof by cases" for the obvious reasons. Verify this tautological implication! □

6.3. Proofs and Theorems

The job of a mathematical proof is to unfailingly preserve truth **in all its steps** as it is developed.

The syntax (SHAPE!) of proofs:

A proof is a finite sequence of formulas —it is our "mathematical argument"— where *EACH formula we write down*, ONE per line with a short explanation to the right, is either

1. an "assumption, also called a hypothesis*" OR an *axiom*,

   OR

2. is obtained from formulas we wrote earlier *IN THIS PROOF* employing *some valid rule*.

   Rules are introduced below!

*"Hypothesis" to be explained on p.278.

Am I allowed in step 1. above to write *an already proved theorem* A?

Yes, because doing so is equivalent to lengthening the proof by adding —instead of just A— ALL OF $\boxed{\ldots, A}$, that is, the *entire proof of* A obtained from axioms only, *not invoking other theorems*.

Programming analogy: I am allowed to invoke **macros** in a program because this is equivalent to writing down explicitly the macro-expansion code.

What are our axioms, our starting assumptions, when we do proofs?

We have two types:

1. Axioms needed by **Logic** (*Logical Axioms*) that are common in all proof-work that we do in *mathematics* or *computer science*.

   ▶ For example, such is the "identity" axiom $x = x$ and the tautology $\neg A \vee A$.

   Both these configurations —"$x = x$" and "$\neg A \vee A$"— define *infinitely many axioms* as their "instances".

   The first allows us to use *ANY* object variable in place of "x" the second allows to use any "statement" (*formula*) in place of A.

2. Axioms needed **to do MATH** in some theory (*Mathematical axioms*).

Here is a *sample* of axioms from a few *MATH theories*:

(i) i. Number theory ("Peano arithmetic") for $\mathbb{N}$:

- $x < y \vee x = y \vee x > y$ (*trichotomy*)
- $\neg x < 0$ this axiom indicates that 0 is *minimal* in $\mathbb{N}$.
- Many others that we omit.

ii. Euclidean Geometry:

- From two distinct points passes *one and only one* line.
- ("Axiom of parallels") From a point A off a line named k —both A and k being on the same plane— passes a unique line on said plane that is parallel to k.
- Many others that we omit.

iii. Axiomatic Set Theory:

- For any set A, we have

$$(\exists y)y \in A \to (\exists x)\Big(x \in A \wedge \neg(\exists z \in A)z \in x\Big)$$

This is the so-called axiom of "foundation" from which one can prove things like $A \in A$ is always *false*.

This axiom incarnates Principles 0-2 in an axiomatic set theory like "ZFC".

It says that *IF* $A \neq \emptyset$ —this is "$(\exists y)y \in A$"— *THEN* there is some element in A —this is the part "$(\exists x)\Big(x \in A$"— *which contains no element of* A —this is the part "$\neg(\exists z \in A)z \in x$".

- And a few others —including the Axiom of Choice, acronym "AC"— that we omit. □

Foundation above tells us, among other things, that we cannot contain all members of a chain

$$\ldots \in x'' \in x' \in x$$

in a set A.

And then we have "*hypotheses*" or "*assumptions*".

Are those not just axioms of logic or math? Not necessarily!

You recall that to prove $A \subseteq B$ you go like this:

"**Let** $x \in A$ for some fixed x". This "Let $x \in A$" is a hypothesis from which you will prove (hopefully) $x \in B$.

It is *NOT* an axiom of logic nor one of mathematics!

6.3.1 Definition. (The *SHAPE* of Logical Axioms)

1. All tautologies; these need no defence as "start-up truths".
2. Formulas of the form $(\forall x)A[x] \to A[t]$, for any formula A, variable x and "object" t.

Notation $A[x]$ denotes our interest in the (potentially) input variable x.

I said "potentially"!

Having written $A[x]$ any notation "$A[t]$" that follows that fact denotes that t has being "**read into**" (or **substituted into**) the input variable x.

▶ x may well be input in A but it is **DEFINITELY NOT** input in $(\forall x)A$.

This t-object can be as simple as an (object) variable y (might be same as x), constant c, or as complex as a "*function call*", $f\Big(g\big(y, h(z)\big), a, b, w\Big)$ where f accepts 4 inputs, g accepts 2 and h accepts one. y, z, w are variables while a and b —**by notational convention**— are unspecified constants.

Mar. 25, 2024

The axiom is true in any theory as it "says" "if A is true for all (values of) x, then it is also true for the specific value t".

The axiom works ONLY IF we take care that **no input variable of t** (say "z") **lands in the scope** of a $(\forall z)$ or a $(\exists z)$ that are embedded in formula A.

If that happens, we say that the free variable z of t was **captured** and we **disallow** this substitution **as illegal**.

So, e.g., If $A[y]$ is $(\exists z)z \neq y$ —which says that for any y-value there is an z-value that is different— *we* cannot *take t to be z* and do $A[z]$. If we do, we get $(\exists z)z \neq z$.

This is false in all domains while the original is true, for example, in the domain of $\mathbb{N}$!

As noted already, "$[x]$" indicates **the free variable of interest to us**. It does not imply that x actually occurs free in A nor does it imply that there may not be *other* free variables in A.

How do I indicate that x, y, z are precisely all the free variables ("inputs") of A? $A(x, y, z)$.

3. Formulas of the form $A[x] \to (\forall x)A[x]$, **for any formula A where the variable x does not occur free in it.**

We wrote "$A[x]$" to speak of our interest in x even though we know (our assumption) that x is non-input in A.

That is, the truth value of A is independent of the value of x and writing —or not writing— "$(\forall x)$" up in front makes no difference.

For example say A is $3 = 3$. This axiom says then, "if $3 = 3$ is true, then so is $(\forall x)3 = 3$".

Sure! $3 = 3$ does **NOT depend** on x. So saying "for all values of x we have $3 = 3$" is the same as saying just "we have $3 = 3$".

4. $x = x$ is the *identity* axiom, no matter what "x" I use to express it. So, $y = y$ and $w = w$ are also instances of the axiom.

5. $x = y \to y = x$ and $x = y \wedge y = z \to x = z$ are the *equality* axioms. They can be expressed equally well using variables other than x and y (e.g., u, v and w).

□

6.3.2 Remark. (The "∃") The symbol ∃ is an abbreviation:

For any formula A, $(\exists x)A[x]$ stands for or is short for $\neg(\forall x)\neg A[x]$.

We also get the tautology (hence theorem)

$$\vdash \overbrace{(\exists x)A}^{\text{using abbrev. of rhs}} \equiv \overbrace{\neg}^{it\ is\ not\ true\ that} \overbrace{(\forall x)\neg A}^{all\ x\ make\ A\ false}$$

This is a **DEFINITION** (a "naming" [of $\neg(\forall x)\neg A$]) **NOT** an axiom!

□ 

The "rules of proving", or rules of inference. These are two up in front —you will find I am grossly miscounting:

6.3.3 Definition. (Rules of Inference)

The rules used in proofs are called *rules of inference* and are these two (actually the second contains infinitely many rules).

1. From $A[x]$ I may infer $(\forall x)A[x]$. Logicians write the up-in-front (also called "primary") rules as fractions without words:

$$\frac{A[x]}{(\forall x)A[x]} \tag{1}$$

this rule we call *generalisation*, or *Gen* in short.

2. I may *construct* (and use) using any tautological implication *that I have verified*, say, this one

$$A_1, A_2, \ldots, A_n \models_{taut} B \tag{2}$$

the rule

$$\frac{A_1, A_2, \ldots, A_n}{B}$$

Example. Seeing readily that $A, A \to B \models_{taut} B$, we have the rule

$$\frac{A, A \to B}{B}$$

This is a very popular rule, known as *modus ponens*, for short *MP*.

 **Worth Saying**. So rules preserve truth. 

Read a rule such as (1) or (2) as saying

If you *already* wrote *all* the formulas of the "numerator" (*in any order*) in a proof, then it is *legitimate to write thereafter in the proof* the denominator formula (of the rule).

We call the numerator *input* or *hypotheses* of the rule and call the denominator *result* or *conclusion*.

□

6.3.4 Remark.

1. The second "rule" above is a rule constructor.

   Any tautological implication we come up with is fair game:

   It leads to a *valid rule* since the name of the game (in a proof) is *preservation/propagation of truth.*

   This is NOT an invitation to learn and memorise infinitely many rules (!) but is rather a license to build your own rules as you go, *as long as you bothered to verify* the **validity of the tautological implication they are derived from**.

2. Gen is a rule that indeed propagates truth: If $A[x]$ is true, that *means* that it is so for all values of x —and all values of any other free variables on which A depends but I did not show in the $[\ldots]$ notation.

   But then so is $(\forall x)A[x]$ true, as it *says precisely the same thing*: "$A[x]$ is true, for all values of x and all values of any other free variables on which A depends but I did not show in the $[\ldots]$ notation".

   The only difference between the two notations is that **I added some notational *emphasis* in the second** —$(\forall x)$.

3. **Hmm**. So is $\forall x$ redundant? Yes, but *ONLY as a formula PREFIX*.

   However, in something like this

   $$x = 0 \to (\forall x)x = 0 \tag{1}$$

   over $\mathbb{N}$ it is **NOT** redundant!

   Dropping $\forall$ we totally change the meaning of (1).

   As is, (1) is *not* a true statement. **For example, if the value of the "input x" (the left one!) is 0, then it is false**.

   However dropping $\forall x$, (1) changes to $x = 0 \to x = 0$ which is a tautology; *always true*.

□ 

6.3.5 Definition. (Theorems)

A theorem is a formula that **appears** at the end of a proof.

Often one writes $\vdash A$ to symbolically say that A is a theorem. If we must indicate that we worked in some specific theory, say ZFC (set theory), then we may indicate this as

$$\vdash_{ZFC} A$$

If moreover we have had some "*non-axiom* hypotheses" (see box on p.278) that form a set Σ, then we may indicate so by writing

$$\Sigma \vdash_{ZFC} A$$

□

Why write Σ —and not Q, R, or C?— for a set of (*non-axiom*) assumptions? Because we reserve upper case latin letters for *SINGLE* formulas. For *sets* of formulas we use *distinguishable* capital letters, so, we chose here a distinguishable Greek capital letters, such as $\Gamma, \Sigma, \Delta, \Phi, \Theta, \Psi, \Omega$. Obviously, Greek capital letters like A, B, E, Z will not do!

6.3.6 Remark. (Hilbert-style proofs) The proof concept as defined is known as a "Hilbert-style proof".

We write them *vertically*, ONE formula per line, every formula consecutively numbered, with annotation to the right of each formula written (this is the "why did I write this?").

Like this

1)	F_1	⟨because⟩
2)	F_2	⟨because⟩
⋮	⋮	⋮
n)	F_n	⟨because⟩

□ 

6.4. Proof Examples

6.4.1 Example. (New (derived) rules) A **derived rule** is one we were not given **up in front** —in 6.3.3— to bootstrap logic, but we can still prove that they propagate truth.

1. We have a new (derived) rule: $(\forall x)A[x] \vdash A[t]$.

   This is called *Specialisation*, or *Spec* **Rule**.

   Aha! We used a *non-axiom hypothesis* here!

   I write a Hilbert proof to show that $A[t]$ is a theorem if $(\forall x)A[x]$ is a (non-axiom) hypothesis (assumption) —shortened to "hyp".

1)	$(\forall x)A[x]$	$\langle$hyp$\rangle$
2)	$(\forall x)A[x] \to A[t]$	$\langle$axiom$\rangle$
3)	$A[t]$	$\langle 1 + 2 +$ MP$\rangle$

2. Taking t to be x we have $(\forall x)A[x] \vdash A[x]$, simply written as $(\forall x)A \vdash A$.

Mar. 27, 2024

3. The *Dual Spec* derived rule:

$$A[t] \vdash (\exists x)A[x] \tag{1}$$

We prove it below, but **first** I must prove:

$$\vdash A[t] \to (\exists x)A[x] \tag{2}$$

Here it goes

$$\begin{array}{lll} 1) & (\forall x)\overbrace{\neg A[x]}^{B[x]} \to \overbrace{\neg A[t]}^{B[t]} & \langle \text{axiom} \rangle \\ 2) & A[t] \to \neg(\forall x)\neg A[x] & \langle 1 + \text{Taut. Impl. (\textbf{contrapositive})} \rangle \\ 2') & A[t] \to (\exists x)A[x] & \langle 2 + \text{using abbreviation ``}\exists\text{''} \rangle \end{array}$$

In step two I used the tautological implication $A \to B \models_{taut} \neg B \to \neg A$. The two sides of "$\models_{taut}$" are called "contrapositives" of each other.

Now, **Dual Spec**:

$$\begin{array}{lll} 1) & A[t] & \langle \text{hyp} \rangle \\ 2) & A[t] \to (\exists x)A[x] & \langle \text{proved above; we quoted a theorem!!} \rangle \\ 3) & (\exists x)A[x] & \langle 1 + 2 + \text{MP} \rangle \end{array}$$

Taking t to be x we have $A[x] \vdash (\exists x)A[x]$, simply written as $A \vdash (\exists x)A$.

□

There are two principles of proof that we state without proving their validity (see [Tou03a, Tou08] if curious).

6.4.2 Remark. (Deduction Theorem and Proof by Contradiction)

1. The *deduction theorem* (also known as "proof by assuming the antecedent" —acronym we use: "**DThm**") states, if

$$\Gamma, A \vdash B \tag{1}$$

then also $\Gamma \vdash A \to B$, **provided** that in the proof of (1), all free variables that **appear in** A were treated as constants (as we say, were "frozen") **AT or BELOW** the point in the proof where A was **inserted as a hypothesis**:

> This "freezing" applies to ALL formulas, B, not just to A in the **entire proof segment below the spot where we said "A is a hypothesis". We cannot apply $\forall$ nor the (derived) operation of assigning a value to such free variables no matter which formula B they occur in.**

6.4.3 Example. ("Everyday" DThm application)

To show $A \subseteq B$ we do $x \in A \to x \in B$ for all x.

To do the latter we pick a fixed ("frozen"!) undisclosed x and assume $x \in A$.

Aha! "**FROZEN**"!

So it **behaves as a constant**. I cannot do $\forall$ to such variables!

Then we proceed to show $x \in B$ for that **same, frozen** x.

Hey! This is an application of the DThm! □

The notation "Γ, A" is standard for the more elaborate $\Gamma \cup \{A\}$.

In practice, this principle is applied to **prove** $\Gamma \vdash A \to B$, **by doing instead** the "easier" (1).

Why "easier"?

(1) We are helped by an *extra hypothesis*, A, and
(2) the formula to prove, B, is *less complex* than $A \to B$.

2. **Proof by contradiction**. To prove $\Gamma \vdash A$ —where A has *no free variables* or, as we say, is *closed* or is a *sentence*— is equivalent to proving the **"constant formula" f** from hypothesis $\Gamma, \neg A$. □ 

6.4.4 Remark. (Ping-Pong) For any formulas A and B, the formula —where I am using way more brackets than I have to, ironically, to *improve* readability—

$$(A \equiv B) \equiv \Big((A \to B) \wedge (B \to A) \Big)$$

is a tautology.

Thus to prove the lhs of the $\equiv$ suffices to prove the rhs.

In turn, to prove the rhs it *suffices* to prove *each* of $A \to B$ and $B \to A$ *separately*. This last idea encapsulates the ping-pong approach to proving equivalences. □

Here are a few applications.

6.4.5 Example. 1. Establish $\vdash (\forall x)(A \wedge B) \equiv (\forall x)A \wedge (\forall x)B$.

By ping-pong.

(I) ($\rightarrow$) Prove $\vdash (\forall x)(A \wedge B) \rightarrow (\forall x)A \wedge (\forall x)B$. By DThm suffices to do $(\forall x)(A \wedge B) \vdash (\forall x)A \wedge (\forall x)B$ instead.

1)	$(\forall x)(A \wedge B)$	$\langle$hyp$\rangle$
2)	$A \wedge B$	$\langle 1 +$ Spec$\rangle$
3)	A	$\langle 2 +$ tautological implication$\rangle$
4)	B	$\langle 2 +$ tautological implication$\rangle$
5)	$(\forall x)A$	$\langle 3 +$ Gen; OK: x is not free in line 1$\rangle$
6)	$(\forall x)B$	$\langle 4 +$ Gen; OK: x is not free in line 1$\rangle$
7)	$(\forall x)A \wedge (\forall x)B$	$\langle 5 + 6 +$ tautological implication$\rangle$

Why the note "OK: x is not free in line 1"?

Because I applied DThm and moved $(\forall x)(A \wedge B)$ to the left of "$\vdash$"(I made it "hyp").

DThm *requires ALL* **FREE** variables of this formula to be *frozen* from the point of insertion down.

In particular I am *NOT allowed* to invoke $(\forall x)$ **IF** x **is free in the hyp line**. **Luckily it is NOT!**

(II) ($\leftarrow$) Prove $\vdash (\forall x)A \wedge (\forall x)B \rightarrow (\forall x)(A \wedge B)$. By DThm suffices to do $(\forall x)A \wedge (\forall x)B \vdash (\forall x)(A \wedge B)$ instead.

1)	$(\forall x)A \wedge (\forall x)B$	$\langle$hyp$\rangle$
2)	$(\forall x)A$	$\langle 1$ + tautological implication$\rangle$
3)	$(\forall x)B$	$\langle 1$ + tautological implication$\rangle$

Complete the above proof!

2. Prove $\vdash (\forall x)(\forall y)A \equiv (\forall y)(\forall x)A$.

   By ping-pong.

   (a) Prove $\vdash (\forall x)(\forall y)A \to (\forall y)(\forall x)A$.
   By DThm suffices to do $(\forall x)(\forall y)A \vdash (\forall y)(\forall x)A$ instead.

1)	$(\forall x)(\forall y)A$	$\langle$hyp$\rangle$
2)	$(\forall y)A$	$\langle 1$ + Spec$\rangle$
3)	A	$\langle 2$ + Spec$\rangle$
4)	$(\forall x)A$	$\langle 3$ + Gen; OK, no free x in line 1$\rangle$
5)	$(\forall y)(\forall x)A$	$\langle 4$ + Gen; OK, no free y in line 1$\rangle$

   (b) Prove $\vdash (\forall y)(\forall x)A \to (\forall x)(\forall y)A$.
   Exercise! □

6.4.6 Exercise. Prove for any A and B — where x is not free in A— that $\vdash (\forall x)(A \to B) \to (A \to (\forall x)B)$. □

6.4.7 Exercise. Prove for any A and B — where x is not free in A— that $A \to B \vdash A \to (\forall x)B$. □

We have seen how to *add* an $(\exists x)$ in front of a formula (6.4.1 3).

How about *removing* an $(\exists x)$-prefix? This is much more complex than removing a $(\forall x)$-prefix:

The the technique of eliminating $\exists$*-prefixes* USES the deduction theorem **in its proof of correctness**.

Technique of removing an $\exists$-prefix: Suppose I have that $(\exists x)A[x]$ is true —either as an **assumption** or a **theorem I proved earlier**— and I want to prove B.

Then I **assume** that —for *some* constant c that does not occur in B— $A[c]$ is true.

In words, "**Let** c be a value (constant!) that makes $A[c]$ true".

That is, I **add** $A[c]$ for an NEW constant c *NOT* in B as a *NEW* non-axiom **hypothesis**.

People annotate this step in a proof as "aux. hyp. related to $(\exists x)A[x]$."

Now I proceed to prove B using all that is known to me —that is, the axioms of the theory $\mathcal{T}$ and the non-axiom hypotheses Γ *and the non-axiom hypothesis* $A[c]$.

I do so by using all free (input-) variables of $A[c]$ as constants in my proof.[b]

[b] This is a side-effect of using the deduction theorem in the proof of correctness of the theorem below that justifies this technique.

Apr. 1, 2024

6.4.8 Metatheorem. (Aux. Hyp. Metatheorem) *Suppose I work within* **theory** $\mathcal{T}$ *and* **hypotheses** Γ *and* I have proved

$$\Gamma \vdash_{\mathcal{T}} (\exists x)A[x]$$

Suppose next I want to get rid of "$(\exists x)$*" by adding the* **hypothesis** $A[c]$ *to* Γ*, where* c *is a* NEW *constant that is not part of* B*,* **nor of my axioms and hypotheses***, and I manage to prove*

$$\Gamma, A[c] \vdash_{\mathcal{T}} B \tag{1}$$

where

(1) *All free variables of* $A[c]$ *were* **frozen** *throughout the proof*

(2) c **does not appear***, as Noted above, in the formulas of* Γ *or in the MATH axioms of* $\mathcal{T}$*.*

Then obtaining (1) under all stated **restrictions** *constitutes a proof of* $\Gamma \vdash_{\mathcal{T}} B$*.*

Intuitively $A(c)$ says "**for SOME** c**,** $A(c)$ **is true**"
Same as $(\exists x)A(x)$: "**for SOME** x**,** $A(x)$ **is true**"**'.**

BUT, Technically, $(\exists x)A(x)$ does *NOT* imply $A(c)$. **For one thing**, you **cannot put your finger on WHAT** c **is!**

For another, you introduce $A(c)$ as a *HYPOTHESIS*: "**Let** c be one that makes $A(c)$ tick".

See Exercises 6.4.11 and 6.4.12.

The "big deal" in Metatheorem 6.4.8 is that normally if you add a **hypothesis** X to the hypotheses Γ and prove

$$\Gamma, X \vdash_{\mathcal{T}} B$$

then you cannot in general get rid of the dependence of the theorem B **on the added hypothesis** X.

Not so with the technique of Metatheorem 6.4.8: You get

$$\Gamma \vdash_{\mathcal{T}} B$$

as if you never *assumed* or *used* $A[c]$!

That is why they call it "auxiliary hypothesis". Once it helps you to prove B it *drops out*; it does not stay around to get credit!

6.4.9 Example. Prove $\vdash (\exists y)(\forall x)A[x,y] \to (\forall x)(\exists y)A[x,y]$.

By the DThm it suffices to prove $(\exists y)(\forall x)A[x,y] \vdash (\forall x)(\exists y)A[x,y]$ instead.

1)	$(\exists y)(\forall x)A[x,y]$	⟨hyp via DThm⟩
2)	$(\forall x)A[x,c]$	⟨aux. hyp. related to 1; for constant c not in the conclusion⟩
3)	$A[x,c]$	⟨2 + Spec⟩
4)	$(\exists y)A[x,y]$	⟨3 + Dual Spec⟩
5)	$(\forall x)(\exists y)A[x,y]$	⟨4 + Gen; OK, no free x in lines 1(DThm) and 2(aux. hyp)⟩

Worth Noting: The "Γ" here is $\{(\exists y)(\forall x)A[x,y]\}$ thus we *do have* $\Gamma \vdash (\exists y)(\forall x)A[x,y]$[b] as required by Theorem 6.4.8.

[b] What I am invoking here is the trivial $X \vdash X$ that is verified by the 1-line proof "1) X ⟨*hyp*⟩".

□

6.4.10 Example. Can I also prove the converse of the above? That is

$$\vdash (\forall x)(\exists y)A[x,y] \to (\exists y)(\forall x)A[x,y] \tag{1}$$

Worth trying.

By the DThm it suffices to prove $(\forall x)(\exists y)A[x,y] \vdash (\exists y)(\forall x)A[x,y]$ instead.

1)	$(\forall x)(\exists y)A[x,y]$	⟨hyp via DThm⟩
2)	$(\exists y)A[x,y]$	⟨1 + Spec⟩
3)	$A[x,c]$	⟨aux. hyp. for 2; NEW c not in the conclusion⟩
4)	$(\forall x)A[x,c]$	⟨3 + Gen; Stop! Forbidden! Illegal "$(\forall x)$": I should treat the free x of aux. hyp. on line 3 as a constant!⟩

Still, can anyone PROVE (1); even if I cannot?

A question like this, *if you are to answer "NO"*, must be resolved by offering a **counterexample**.

That is, a SPECIAL, SIMPLE case of A for which I can clearly see that the claim is **false**.

Here is one such (counter)example over the set $\mathbb{N}$:

$$\underbrace{(\forall x)(\exists y)\overbrace{x=y}^{\text{"the } A\text{"}}}_{\mathbf{t}} \to \underbrace{(\exists y)(\forall x)\overbrace{x=y}^{\text{"the } A\text{"}}}_{\mathbf{f}} \tag{1}$$

□

Here is another non-theorem. We have the **axiom** $A \to (\forall x)A$ if x is not free in A. **Can we relax the restriction on x?**

No. If we had $\vdash A \to (\forall x)A$ with no restrictions then look at the **special case**

$$x = 0 \to (\forall x)x = 0 \tag{2}$$

on $\mathbb{N}$.

We already saw that this is NOT true for all x —**not a theorem then!**

In fact over $\mathbb{N}$, (2) is false if the in input x is 0: $\overbrace{0 = 0}^{\mathbf{t}} \to \overbrace{(\forall x)x = 0}^{\mathbf{f}}$.

6.4.11 Exercise. (Important "confusion remover") One might be confused by the act of *adding the hypothesis* $A(c)$ whenever we have $(\exists x)A(x)$.

Some lapse of judgement might construe this as an implication:

$$(\exists x)A(x) \to A(c) \tag{1}$$

The above is false!! NOT a theorem!!

Working over the natural numbers, Prove by finding a very simple $A(x)$ and a specific appropriate constant c that (1) fails for this A and c so it is NOT a theorem! □

6.4.12 Exercise. (Important "confusion remover" #2) Prove by an *EASY* counterexample that $(\exists x)A[x] \to A[x]$ is not provable either. □

Another useful principle that **can** be proved, but we will not do so, is that one can *replace equivalents-by-equivalents.* That is, if C is some formula, and if I have

1. Let $A \equiv B$, **via proof**, or **via assumption**, and also
2. A is a subformula of C

then I can **replace one (or more) occurrence(s) of** A in C (as subformula(s)) by B and call the resulting formula C'.

I will be guaranteed the theorem $C \equiv C'$.

That is, from $A \equiv B$, I can prove $C \equiv C'$.

This principle is called the *equivalence theorem*.

Let's do a couple of ad hoc additional examples before we move to the section on Induction.

6.4.13 Example. $A \to B \vdash (\forall x)A \to (\forall x)B$.

By the DThm it suffices to prove $A \to B, (\forall x)A \vdash (\forall x)B$ instead.

1) $A \to B$ $\langle$hyp$\rangle$
2) $(\forall x)A$ $\langle$hyp from DThm$\rangle$
3) A $\langle 2 +$ Spec$\rangle$
4) B $\langle 1 + 3 +$ MP$\rangle$
5) $(\forall x)B$ $\langle 4 +$ Gen; OK as the DThm hyp. (line 2) has no free $x\rangle$

We don't CARE whether Line 1 has free x's.

□

6.4.14 Example. (Substitution Theorem) We have $A[x] \vdash A[t]$ for any term t.

Indeed,

1)	$A[x]$	$\langle$hyp$\rangle$
2)	$(\forall x)A[x]$	$\langle 1 + \text{Gen}\rangle$
3)	$A[t]$	$\langle 2 + \text{Spec}\rangle$

□

6.4.15 Example. We have $A \to B \vdash (\exists x)A \to (\exists x)B$.

Proof via DThm, that is, prove

$$A \to B, (\exists x)A \vdash (\exists x)B$$

instead.

1)	$A[x] \to B[x]$	$\langle$hyp$\rangle$
2)	$(\exists x)A[x]$	$\langle$hyp via DThm$\rangle$
3)	$A[c]$	$\langle$aux. hyp. for 2$\rangle$
4)	$A[c] \to B[c]$	$\langle 1 + 6.4.14$; OK no free x in lines #2, 3$\rangle$
5)	$B[c]$	$\langle 3 + 4 + \text{MP}\rangle$
6)	$(\exists x)B[x]$	$\langle 5 + \text{Dual Spec}\rangle$

□

6.4.16 Example. READ ME! Refer to 6.3.2. Let us apply it to $\neg A$ for arbitrary A. We get

$$\vdash (\exists x)\neg A \equiv \neg(\forall x)\neg\neg A \tag{1}$$

Since $A \equiv \neg\neg A$ is a tautology, hence a theorem

Pause. Why "hence a theorem"?◀

we apply the *equivalence theorem* (p.310) and tautological implication[†] and obtain:

$$\vdash \neg(\forall x)A \equiv (\exists x)\neg A \tag{2}$$

Applying another tautological implication to (2) we move the left-most $\neg$ just past the "$\equiv$" and get

$$\vdash (\forall x)A \equiv \neg(\exists x)\neg A$$

which is of the same form as 6.3.2 with the roles of $\exists$ and $\forall$ reversed. □

[†]We have (1) and $\vdash \neg(\forall x)\neg\neg A \equiv \neg(\forall x)A$ as the lhs of the implication.

6.4.17 Example. $A \equiv B \vdash \overbrace{(\forall x)A}^{C} \equiv \overbrace{(\forall x)B}^{C'}$.

True due to the equivalence theorem! "C" is "$(\forall x)A$". We replaced (one occurrence of) A by B in C, and we have assumed as starting point that $A \equiv B$. □

6.4.18 Exercise. Prove $A \equiv B \vdash (\forall x)A \equiv (\forall x)B$ without relying on the equivalence theorem. Rather use 6.4.13 in your proof, remembering the ping-pong tautology (6.4.4). □

6.5. Induction

Apr. 3, 2024

In Remark 4.4.46 we concluded with a formulation —(2) on p.169— of the *minimal condition* (MC) for any order $<$.

See (†) below:

Since we often[†] depict a class $\mathbb{A}$ as $\mathbb{A} = \{x : F[x]\}$ for some "entrance property" $F[x]$, we have

The statement "*some order $<$ has MC*" is captured by the statement

For any "property", that is, formula $F[x]$, we have that the following is true

$$(\exists a)F[a] \to (\exists a)\Big(F[a] \wedge \neg(\exists y)\big(y < a \wedge F[y]\big)\Big) \qquad (\dagger)$$

In what follows **in this Section Only** the specific "$<$" we are using is the total order **on** $\mathbb{N}$, hence it (*satisfies trichotomy*) and thus the concepts *minimal* and *minimum* coincide as we know (4.4.43).

[†] "Often", not "always". There are more classes —in fact, there are even *more SETS*— than formulas ("properties").

We will prove that MC (†) is equivalent to the *Principle* of so-called "*Strong Induction*" or "*Course-of-Values Induction*" —CVI— on $\mathbb{N}$.

In the proof I will use the equivalence theorem and three easy THEOREMS (they are tautologies, hence axioms, **hence theorems**[†]) as well as 6.3.2.

Theorem 1. $\vdash \neg A \vee B \equiv A \to B$

Theorem 2. $A \to B \equiv \neg B \to \neg A$ (contrapositive).

Theorem 3. $\neg(A \vee B) \equiv \neg A \wedge \neg B$ and also $\neg(A \wedge B) \equiv \neg A \vee \neg B$.

These are the well-known "de Morgan" equivalences.[‡]. The first intuitively says "$A \vee B$ is false iff *both* A and B are false". The second intuitively says "$A \wedge B$ is false iff at least one of A and B is false".

READ ME! (The proof below will Not be examined

Our proof below is written as a *conjunctional* $\Leftrightarrow$-chain, written vertically with annotation "$\langle \ldots \rangle$" to the right of *each* $\Leftrightarrow$.

The conjunctional $\Leftrightarrow$ (and $\Rightarrow$) were introduced on p.46

Such proofs are called *Equational* ([DS90, GS94, Tou08]).

[†]Let A be an axiom. Then

1) A ⟨axiom⟩

is a 1-line proof, hence A is a theorem.

[‡]Often called "de Morgan Laws".

So let $P[x]$ be an arbitrary "property" (formula!) of the variable x.

We start our "Equational proof" with (†) at the top of the chain, but where we replaced the arbitrary "F" there with "$\neg P$" here.

$$(\exists a)\neg P[a] \to (\exists a)\Big(\neg P[a] \wedge \neg(\exists y)\big(y < a \wedge \neg P[y]\big)\Big)$$

$\Leftrightarrow$ ⟨using 6.3.2 and equiv. thm (p.310) removing double negations⟩

$$\neg(\forall a)P[a] \to \neg(\forall a)\neg\Big(\neg P[a] \wedge (\forall y)\neg\big(y < a \wedge \neg P[y]\big)\Big)$$

$\Leftrightarrow$ ⟨contrapositive⟩

$$(\forall a)\neg\Big(\neg P[a] \wedge (\forall y)\neg\big(y < a \wedge \neg P[y]\big)\Big) \to (\forall a)P[a]$$

$\Leftrightarrow$ ⟨two applications of de Morgan and equiv. thm⟩

$$(\forall a)\Big(P[a] \vee \neg(\forall y)\big(\neg y < a \vee P[y]\big)\Big) \to (\forall a)P[a]$$

$\Leftrightarrow$ ⟨Theorem 1 above + equiv. thm (twice)⟩

$$(\forall a)\Big(\underbrace{(\forall y)\big(y < a \to P[y]\big) \to P[a]}_{\text{I.H.}}\Big) \to (\forall a)P[a] \quad (\ddagger)$$

Fix a. Then if **for all $y < a$, $P[y]$ is true** (I.H.) **implies** $P[a]$ (I.S.)

We have proved that (†) —the least principle for $\mathbb{N}$— of the previous page is **equivalent** to the formula on line (‡).

(‡) embodies the *Strong Induction* or *Course-of-Values Induction* (CVI) Principle:

> To prove $(\forall a)P[a]$, where P is a property over $\mathbb{N}$ and a is a $\mathbb{N}$-variable, it suffices to do **TWO steps**:
>
> 1. Fix —but do not disclose— an a (a is *arbitrary* but *fixed*).
> 2. From the (blue) hypothesis about our fixed unspecified a — underlined on line (‡) in the formula display above that is called *Induction Hypothesis* or *I.H.*— that $P[y]$ is true for all $y < a$, I must prove that so is $P[a]$.
>
>   This "must prove" step is called the *Induction Step*, or *I.S.*

Hmm! All inductions have a "Basis", typically at 0. Doesn't this one? It does!

Well, for $a = 0$,

1. The "blue implication" in line (‡) of the previous page reads

$$\textbf{If } \underbrace{(\forall y)\big(\overbrace{y < 0}^{\textbf{f}} \to P[y]\big)}_{I.H.\ \textbf{t}} \textbf{ then } \underbrace{P[0]}_{\textit{prove } \textbf{t directly}}$$

2. The "I.H." part above **IS true**, but gives no information or help to prove the truth of $P[0]$. So the proof of $P[0]$ must be done **directly and unassisted by an I.H. This is our familiar from other courses (e.g., Calculus) "Basis"!**

There is another simpler induction principle that we call, well, "*simple* induction":

$$\frac{P[0], P[x] \to P[x+1]}{P[x]} \qquad (SI)$$

"(SI)" for Simple Induction. That is, to prove $P[x]$ for all x (denominator) do *three* things:

Step 1. Prove/verify $P[0]$

Step 2. **Assume** $P[x]$ for fixed ("frozen") x (unspecified!).

Step 3. **prove** $P[x+1]$ for that same (previously frozen) x.

The assumption is the I.H. for simple induction.

The I.S. is Step 3 that proves $P[x+1]$.

Note that what is described here is precisely an application of the Deduction theorem towards proving "$P[x] \to P[x+1]$", that is, **proving the implication for every given** x.

Step 4. If you have done **Step** 1. through **Step** 3. above, then you **announce that you have proved** $P[x]$ (for all x is implied!)

THIS PART (6.5.1 and 6.52) is NOT Examinable.

Is the principle (SI) *correct*? I.e., if I do all that the numerator of (SI) asks me to do (*equivalently*, **Steps** 1. – 3.), then do I *really* get that the denominator is true (for all x implied)? *YES!*

6.5.1 Theorem. (*MC* $\rightarrow$ *SI*; **Skip Proof**) *The validity of (SI) is a consequence of MC (least principle) on* $\mathbb{N}$.

Proof. Suppose (SI) is *not* correct.

Then, for some property $P[x]$, *despite having completed* **Steps** *1. – 3.,* $P[x]$ *is* not true *for all* x*!*

Then,

$$\boxed{\text{let } n \in \mathbb{N} \text{ be } \textit{smallest} \text{ such that } P[n] \text{ is } \textit{false}.}$$

Now, $n > 0$ since I *did* verify the truth of $P[0]$ (**Step** 1.).

Thus, $n - 1 \geq 0$.

But then, when I *proved* "$P[x] \rightarrow P[x+1]$ for all x (in $\mathbb{N}$)" —in **Steps** 2. and 3.— this includes **proving**

$$P[n-1] \rightarrow \underbrace{P\left[\overbrace{n}^{smallest}\right]}_{false} \tag{4}$$

By the smallest-ness of n, $P[n-1]$ is *true*, hence $P[n]$ is true after all, by (4).

 I have just contradicted that $P[n]$ is false! 

(SI) works if MC does! □

In fact, MC and SI are equivalent principles.

6.5.2 Theorem. (*SI* $\rightarrow$ *MC*; **Skip Proof**) *Conversely to the previous theorem (6.5.1), if* SI *on* $\mathbb{N}$ *works, then* $\mathbb{N}$ *has MC.*

Proof. By contradiction, I assume I have SI, but that *MC fails*.

So, there is a nonempty $S \subseteq \mathbb{N}$ that has no least element.

I will get a contradiction by showing that $\overline{S} \stackrel{Def}{=} \mathbb{N} - S$ is all of $\mathbb{N}$ (hence $S = \emptyset$).

I apply *SI* to the property

$$P(x) \stackrel{Def}{\equiv} \{0, 1, \ldots, x\} \subseteq \overline{S}$$

1. Basis. $P(0)$ says $\{0\} \subseteq \overline{S}$ which is equivalent to $0 \in \overline{S}$; true since if $0 \in S$ that would contradict assumption on S.
2. Fix x and assume (I.H.) $P(x)$ —i.e., $\{0, 1, \ldots, x\} \subseteq \overline{S}$.
3. $P(x+1)$ says $\{0, 1, \ldots, x, x+1\} \subseteq \overline{S}$. To prove this, note:

By 2., we have $\{0, 1, \ldots, x\} \subseteq \overline{S}$ so if $x + 1 \in S$ instead, then it would be smallest in S, contradicting hypothesis about S.

Thus I MUST have also $\{0, 1, \ldots, x, x + 1\} \subseteq \overline{S}$ —and hence $P(x + 1)$ is true.

By SI, I have $P(x)$ true for all x, thus $\{0, 1, \ldots, x\} \subseteq \overline{S}$ for all x. In particular, $x \in \overline{S}$ for all x

But then $S = \emptyset$. A contradiction! □

Since we have CVI equivalent to MC we now have

JUST KNOW THIS Corollary; NOT its proof.

6.5.3 Corollary. **All three of CVI, SI and MC** *are equivalent principles over* $\mathbb{N}$.

6.6. Induction Practice

To begin with, there are "properties" to prove that are only valid for all $n \geq k$ for some constant $k > 0$.

This is the domain where we have to stay in during the proof.

Thus for those the I.H. MUST "pick a fixed unspecified $n \geq k$".

The points $n = 0, 1, \ldots, k-1$ are outside the domain so are "illegal".

Thus the "*Basis*" (same as "Beginning") of the induction must be for $n = k$.

As an example, the smallest n where $n + 3 < 2^n$ is true is $n = 3$ (verify!).

We can prove by induction

$$n + 3 < 2^n, \text{ for } n \geq 3$$

verifying as Basis the case $n = 3$.

Another example:

The statement "n has a prime factor" is *erratic* for $n < 2$.

For $n = 1$ it is *false* and for $n = 0$ it is *true* (every number is a factor of zero).

So one must take as domain of truth of the quoted blue property the set $\{n \in \mathbb{N} : n \geq 2\}$. 2 is the Basis —the Beginning.

6.6.1 Example. This is the "classic first example of induction use" in the discrete math bibliography! Prove that

$$0+1+2+\ldots+n=\frac{n(n+1)}{2} \tag{1}$$

So, the property to prove is the statement (1).

One must learn to not have to rename the various "properties" that we encounter as "$P[n]$".

I will use SI. So let us do the *Basis*. Boundary case is $n=0$. We verify: $lhs=0$. $rhs=(0\times 1)/2=0$. Good!

Fix n and take the expression (1) as I.H. (WHY "FIX n"? See (SI) on p.319).

Do the I.S. Prove:

$$0+1+2+\ldots+n+(n+1)=\frac{(n+1)(n+2)}{2}$$

Here it goes

$$\begin{aligned}0+1+2+\ldots+n+(n+1) &\overset{\text{using I.H.}}{=} \frac{n(n+1)}{2}+(n+1)\\ &=(n+1)(n/2+1)\\ &=\frac{(n+1)(n+2)}{2}\end{aligned}$$

□

I will write more concisely in the examples that follow.

Apr. 5, 2024

6.6.2 Example. Same as above but doing away with the "0+". Again, I use SI.

$$1+2+\ldots+n=\frac{n(n+1)}{2} \tag{1}$$

- *Basis.* $n=1$: (1) becomes $1=(1\times 2)/2$. True.
- Take (1) as I.H. with fixed n.
- I.S.:

$$\begin{aligned} 1+2+\ldots+n+(n+1) &\overset{\text{using I.H.}}{=} \frac{n(n+1)}{2}+(n+1) \\ &= (n+1)(n/2+1) \\ &= \frac{(n+1)(n+2)}{2} \end{aligned}$$

□

6.6.3 Example. Prove

$$1 + 2 + 2^2 + \ldots + 2^n = 2^{n+1} - 1 \tag{1}$$

By SI.

- Basis. $n = 0$. $lhs = 1 = 2^0 = 2^1 - 1 = rhs$. True.
- As I.H. **take (1) for fixed** n**.**
- I.S.

$$\begin{aligned} 1 + 2 + 2^2 + \ldots + 2^n + 2^{n+1} &\overset{\text{using I.H.}}{=} 2^{n+1} - 1 + 2^{n+1} \\ &= 2 \cdot 2^{n+1} - 1 \\ &= 2^{n+2} - 1 \end{aligned}$$

□

6.6.4 Example. (Euclid) Every natural number $n \geq 2$ has a prime factor.

> p is prime iff
>
> (i) $p > 1$
>
> AND
>
> (ii) The only divisors of p are 1 and p

I do CVI (as you will see why!)

- *Basis*: For $n = 2$ we are done since 2 is a prime and $2 = 2 \times 1$.[†]
- I.H. **Fix an n and assume** the claim for *all* k, such that $2 \leq k < n$.
- I.S.: **Prove for n**: Two subcases:
  1. If n is prime, then OK! n divides n.
  2. If not, then $n = a \times b$, where $a \geq 2$ **and** $b \geq 2$. By I.H. —**from** $a < n$— a has a prime factor, thus so does $n = a \cdot b$. □

> You see? The I.H. in **SI** says that "$n-1$ has a prime factor". But $n-1$ is NOT a factor of n except when $n = 2$. **So this kind of I.H. does not help!**
>
> On the other hand, the I.H. in **CVI ASSUMES** that ALL $2 \leq a < n$ have a prime factor. SO we can proceed as above!

[†]You will recall that a number $\mathbb{N} \ni n > 1$ is a *prime* iff its **only** factors are 1 and n.

6.6.5 Example. Let

$$\begin{aligned} b_1 &= 3,\ b_2 = 6 \\ b_k &= b_{k-1} + b_{k-2},\ \text{for } k \geq 3 \end{aligned}$$

Prove by induction that b_n is divisible by 3 for $n \geq 1$. (Be careful to distinguish between what is *Basis* and what are *Cases* arising from the **induction step**!)

Proof. So the boundary condition is (from the underlined part above) $n = 1$. This is the *Basis*.

1. *Basis*: For $n = 1$, I have $b_1 = 3$ and this is *divisible by 3*. We are good.
2. *I.H.* **Fix an arbitrary** n and **assume claim** for all k such that $1 \leq k < n$ —that is, **assume theorem for all predecessors of** n **down to 1**.
3. *I.S.* **Prove claim** for the above fixed n. There are two cases, as the I.H. is *not useable* for the SMALLEST possible value of FIXED n: $n = 2$. The I.S. **MUST** work for **ANY** "FIXED" unspecified n!

Why I.H. not "usable" for $n = 2$? Because $b_n = b_2$ requires entries b_0 and b_1.

The red entry does not exist since the sequence starts with b_1. So,

Case 1. $n = 2$. DIRECTLY. I am OK as $b_2 = 6$; it *is* divisible by 3.

Case 2. $n > 2$. Is b_n divisible by 3? Well, $b_n = b_{n-1} + b_{n-2}$ **in this case**. By I.H. (valid for all k: $1 \leq k < n$) I have

that $b_{n-1} = 3t$ and $b_{n-2} = 3r$, for some integers t, r. Thus, $b_n = 3(t + r)$. Done! □

Here are a few additional exercises for you to try —please do try!

6.6.6 Exercise.

1. Prove that $2^{2n+1} + 3^{2n+1}$ is divisible by 5 for all $n \geq 0$.
2. Using induction prove that $1^3 + 2^3 + \ldots + n^3 = \left[\frac{n(n+1)}{2}\right]^2$, for $n \geq 1$.
3. Using induction prove that $\sum_{i=1}^{n+1} i2^i = n2^{n+2} + 2$, for $n \geq 0$.
4. Using induction prove that $\sqrt{n} < \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \ldots + \frac{1}{\sqrt{n}}$, for $n \geq 2$.
5. Let

$$\begin{aligned} b_0 &= 1,\ b_1 = 2,\ b_3 = 3 \\ b_k &= b_{k-1} + b_{k-2} + b_{k-3},\ \text{for } k \geq 3 \end{aligned}$$

Prove by induction that $b_n \leq 3^n$ for $n \geq 0$. (Once again, be careful to distinguish between what is *basis* and what are *cases* arising from the **induction step**!) □

Chapter 7

Inductively defined sets; Structural induction

An example of an inductively defined set is the following.

Suppose you want to define *by finite means*, and define *precisely*, the set of all *"simple" arithmetical expressions* that just use the numbers $1, 2, 3$, the operations $+$ and $\times$, and round brackets.

Then you would define:

The set of said *simple arithmetical expressions* is the *smallest* set ($\subseteq$-smallest) that

1. Contains each of $1, 2$ and 3.
2. If it contains expressions E and E', then it also contains $(E + E')$ and $(E \times E')$.

Example: $(1 + (3 \times 2))$

Some folks would add a 3rd ("**fake**") requirement that reads "nothing else is in the set unless so demonstrated using 1. 2. above" and they omit "smallest".

Really?!

How exactly would you "so demonstrate"?

Apr. 8, 2024

Before we get to the general definitions, let us finesse our construction and propose some terminology. First off, we are building a set of simple Arithmetic Formulas.

We do so as follows:

(a) **Simple (Initial) Objects**. First off, in step 1. above we say that 1, 2 and 3 are *the initial objects* of our *recursive*/*inductive* definition.

We throw them IN the set we are building.

(b) **Complex Objects via Rules**. In step 2. we say that *the string* $(E+E')$ is obtained by an *operation* (*on strings* **already in the set we are building**) that is available to us, depicted as a "blackbox" below, which we named "+".

$$\begin{array}{l} E \\ \longrightarrow \\ \longrightarrow \\ E' \end{array} \boxed{\ +\ } \longrightarrow (E+E')$$

In words, the operation *concatenates from left to right the strings*

$$\text{“(”, “}\underbrace{E}_{input1}\text{”, “+”, “}\underbrace{E'}_{input2}\text{”, and “)”}$$

to form the string "$(E+E')$".

Similar comments for the operation "$\times$".

7.0.1 Definition. ("Operations" or "Rules" and "Closed Under") An "**operation or rule on a set** S" is just a function $f : S \to S$. That is, **both** inputs and outputs of f are in S.

Examples.

- Let S be the set of all first-order formulas. An f which, whenever it receives the formulas A and $A \to B$ returns the formula B, is nothing other than the "rule MP" on the set of all formulas.

  Logicians write the rule as $\dfrac{A, A \to B}{B}$ rather than writing it as $f(A, A \to B) = B$ or

  $$A, A \to B \overset{f}{\mapsto} B$$

  This rule is NOT total on S.

- Let again S be the set of all first-order formulas. For each variable symbol x, define the rule g_x that whenever it inputs the formula A it returns the formula $(\forall x)A$.

  This rule is nothing other than the "rule Gen" (with *quantified variable* x) on the set of all formulas.

  Logicians write the rule as $\dfrac{A}{(\forall x)A}$ rather than as $g_x(A) = (\forall x)A$.

  This rule IS total on S.

End Examples

We say that S **is closed under the operation/rule** f.

Operations are often denoted by the generic "O" —for Operation— rather than "f", with primes or subscripts as needed if we have many of those.

Operations $O : S \to S$ are sets since S is, and $O \subseteq S \times S$.

□

7.0.2 Definition. (Closure) Given a set of *initial objects* $\mathcal{I}$ and a *set* of *operations* $\mathcal{O} = \{O_0, O_1, O_2, \ldots\}$.

The object $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ is called *A closure of $\mathcal{I}$ under $\mathcal{O}$* —or *A* set inductively defined by the pair $(\mathcal{I}, \mathcal{O})$.

This $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ **is** the $\subseteq$**-smallest** *set* S that satisfies

1. $\mathcal{I} \subseteq S$.

2. S is *closed under* all *operations in* $\mathcal{O}$, or simply, closed under $\mathcal{O}$ or, even more simply is $\mathcal{O}$-closed.

   The "$\subseteq$**-smallest**" part means: Any set T that satisfies 1. and 2. (just as S does) also satisfies $S \subseteq T$.

The definition does *NOT* guarantee that $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ exists, or that it is unique (hence the indefinite article "*A*"), or, *IF* it does exist, *IF* it is a **set**.

The set $\mathcal{O}$ may be infinite but is countable: That is guaranteed by the indexing $i \mapsto O_i$ being from $\mathbb{N}$ onto $\mathcal{O}$. □

Nice definition, but does $\mathrm{Cl}(\mathcal{I},\mathcal{O})$ *exist* given $\mathcal{I}$ and $\mathcal{O}$? Is it a *set*, as we want/hope?

Yes and Yes.

But first,

7.0.3 Theorem. (Easy) *For any choice of* $\mathcal{I}$ *and* $\mathcal{O}$, if a set $\mathrm{Cl}(\mathcal{I},\mathcal{O})$ exists, *then it is* **unique**.

Proof. Imagine that the definition of $\mathrm{Cl}(\mathcal{I},\mathcal{O})$ ambiguously might lead to (or produces) two sets, S and T —with the properties 1. and 2. of 7.0.2.

Then, letting S pose as closure, we get $S \subseteq T$ from 7.0.2.

Next, letting T pose as closure, we get $T \subseteq S$, again from 7.0.2. Thus $S = T$. □

7.0.4 Theorem. (Know This Theorem; Skip Proof)
For any choice of $\mathcal{I}$ and $\mathcal{O}$ with the restrictions of Definition 7.0.2 the class $\mathrm{Cl}(\mathcal{I},\mathcal{O})$ exists and IS a set.

Proof. We have to check and note a few things.

1. By 4.1.5, for each O_i, $\mathrm{ran}(O_i)$ is indeed a set (because $O_i \subseteq Q \times Q$ for some set Q).
2. The class $\mathbb{F} = \{\mathrm{ran}(O_i) : i = 0, 1, 2 \ldots\}$ is a set. This is so by **Principle** 3, since I can index all members of $\mathbb{F}$ by assigning unique indices from $\mathbb{N}$ to each of its members (and $\mathbb{N}$ is a set by **Principle** 0).
3. By 2. above and 2.5.11, $\bigcup \mathbb{F}$ is a set, and **so is** $T = \mathcal{I} \cup \bigcup \mathbb{F}$.
4. T contains $\mathcal{I}$ as a subset **and** is $\mathcal{O}$-closed since any O_i-output is in $\mathrm{ran}(O_i) \subseteq \bigcup \mathbb{F}$.
5. The family of sets $\mathbb{G} = \{S : \mathcal{I} \subseteq S$ **and** S is $\mathcal{O}$-closed$\}$ contains the **set** T as a member. Thus (cf. 2.5.10)

$$C \stackrel{Def}{=} \left(\bigcap \mathbb{G}\right) \subseteq T$$

is a set.

Since EACH set S in $\mathbb{G}$ contains $\mathcal{I}$ and is $\mathcal{O}$-closed, so is C (*Exercise*!)

But $C \subseteq S$ for all such sets S the way C is defined.

So C is $\subseteq$-smallest among sets in $\mathbb{G}$.

Done. □

7.1. Induction over a closure

7.1.1 Definition. Let a pair $(\mathcal{I}, \mathcal{O})$ be given as above.

We say that a property

$P[x]$ **propagates with** $\mathcal{O}$ *iff*, for each $O_i \in \mathcal{O}$, **whenever all the inputs** x_i **of** O_i satisfy $P[x]$, then **the output value** y **returned by** O_i —for said inputs— satisfies $P[x]$ as well.

We can also say that the property $P[x]$ is preserved by $\mathcal{O}$.

□

7.1.2 Lemma. *For all $(\mathcal{I}, \mathcal{O})$ and a property $P[x]$, if the latter propagates with $\mathcal{O}$, then the class $\mathbb{A} = \{x : P[x]\}$ is closed under $\mathcal{O}$ (is $\mathcal{O}$-closed).*

Proof. So let $O_i \in \mathcal{O}$ have n input variables, $x_1, x_2, \ldots, x_n$. Let the input values chosen —$a_1, \ldots, a_n$— be all in $\mathbb{A}$. Thus

$$P[a_i], \text{ for all } i = 1, \ldots, n$$

By the propagation assumption, if $O_i(a_1, \ldots, a_n) = b$, then $P[b]$ is true, hence the output $b \in \mathbb{A}$. □

7.1.3 Theorem. (Learn the Statement; NOT the proof!)
Let $\mathrm{Cl}(\mathcal{I},\mathcal{O})$ *and a property* $P[x]$ *be given. Suppose we have done the following steps:*

1. *We showed that for* **each** $a \in \mathcal{I}$, $P[a]$ *is true; i.e.,* **all members of** $\mathcal{I}$ **have the property**.
2. *We showed that property* $P[x]$ **propagates** *with* $\mathcal{O}$.

Then **every** $a \in \mathrm{Cl}(\mathcal{I},\mathcal{O})$ *has property* $P[x]$ *—i.e.,* $P[a]$ **is true**.

Naturally, the technique encapsulated by 1. and 2. of 7.1.3 is called "induction over $\mathrm{Cl}(\mathcal{I},\mathcal{O})$" or "**structural induction**" over $\mathrm{Cl}(\mathcal{I},\mathcal{O})$.

▶ Note that —**in Practice**— for each $O_i \in \mathcal{O}$ the "propagation of property $P[x]$" **will be split into**

1. The **I.H.** (assume all the **inputs** of O_i have the property)
2. Followed by the **I.S.** (prove that the **output** of O_i has the property).

Proof. (of 7.1.3) Let us write

$$\mathbb{A} \overset{Def}{=} \{x : P[x]\}$$

Thus, 1. in 7.1.3 translates to

$$\mathcal{I} \subseteq \mathbb{A} \qquad (*)$$

2. in 7.1.3 yields (by Lemma 7.1.2)

$$\mathbb{A} \text{ is } \mathcal{O}\text{-closed} \qquad (**)$$

Now we cannot directly apply Definition 7.0.2 item (3) and say "by $(*)$ and $(**)$ we have"

$$\mathrm{Cl}(\mathcal{I}, \mathcal{O}) \subseteq \mathbb{A}$$

because in 7.0.2 the "*sets* T" that fulfil "1. and 2." must be, well, **sets**; **not (possibly) proper classes**.

Here is the workaround: Get any set that has $\mathcal{I}$ as a subset and is $\mathcal{O}$-closed.

As such we can take the T used in the proof of 7.0.4, or we can take $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ itself!

We take $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ here.

By $(*)$ and $(**)$ Q has $\mathcal{I}$ as subset and is $\mathcal{O}$-closed.

$$Q = \mathrm{Cl}(\mathcal{I}, \mathcal{O}) \cap \mathbb{A} \qquad (***)$$

But Q is a *set* by 2.3.5 and thus

$$\mathrm{Cl}(\mathcal{I}, \mathcal{O}) \overset{7.0.2,3.}{\subseteq} Q \overset{(***)}{\subseteq} \mathbb{A}$$

The last inclusion immediately translates to

$$\boxed{x \in \mathrm{Cl}(\mathcal{I}, \mathcal{O}) \text{ implies } P[x] \text{ is true}} \qquad \square$$

7.1.4 Example. Let $S = \mathrm{Cl}(\mathcal{I}, \mathcal{O})$ where $\mathcal{I} = \{0\}$ and $\mathcal{O}$ contains just one operation,

$$n \longrightarrow \boxed{+1} \longrightarrow n+1 \tag{1}$$

By induction over S, I can show $S \subseteq \mathbb{N}$, or **ALL $x \in S$ are in $\mathbb{N}$**.

▶ The "$P[x]$" is "$x \in \mathbb{N}$"

1. **So $P[0]$ is true. I verified the property for all members of $\mathcal{I}$.**

2. That the property propagates with our operation is captured by (1) above: if $\underbrace{n \in \mathbb{N}}_{prop.\ for\ n}$, then $\underbrace{n+1 \in \mathbb{N}}_{prop.\ for\ n+1}$. **Done!**

Can we also show that $\mathbb{N} \subseteq \mathrm{Cl}(\mathcal{I}, \mathcal{O})$? **Yes**:

In this direction I do SI over $\mathbb{N}$ on variable n.

The property to prove now is "$\overbrace{n \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})}^{\text{"}P[x]\text{"}}$".

For $n = 0$, $n \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})$ since $0 \in \mathcal{I} \subseteq \mathrm{Cl}(\mathcal{I}, \mathcal{O})$ by 7.0.2 (1).

Now, fix n and say $(I.H.)$ $\overbrace{n \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})}^{P[n]}$. Since $\mathrm{Cl}(\mathcal{I}, \mathcal{O})$ is closed under the operation $n \mapsto n+1$, we have $\underbrace{n+1 \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})}_{P[n+1]}$ by 7.0.2.

So, by SI, all $n \in \mathbb{N}$ have the "property $n \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})$". That is,

$$\mathbb{N} \subseteq \mathrm{Cl}(\mathcal{I}, \mathcal{O}) \qquad \square$$

Thus the induction over a closure generalises *SI*.

7.1.5 Example. Let $A = \{a, b\}$.

Let $\mathcal{I} = \{\lambda\}$, let $\mathcal{O}$ consist of **one** operation R:

$$X \longrightarrow \boxed{R} \longrightarrow aXb \tag{3}$$

where "aXb" means concatenation of the strings a, X and b in that order.

We claim that $\mathrm{Cl}(\mathcal{I}, \mathcal{O}) = \{a^n b^n : n \geq 0\}$, where for any string X,

$$X^n \stackrel{Def}{=} \underbrace{XX\ldots X}_{n \text{ copies of } X}$$

If $n = 0$, "0 copies of X" means λ.

Let us write $S = \{a^n b^n : n \geq 0\}$.

1. For $\mathrm{Cl}(\mathcal{I}, \mathcal{O}) \subseteq S$ we do induction over the closure to prove that all $x \in \mathrm{Cl}(\mathcal{I}, \mathcal{O})$ satisfy $x \in S$ ("the property") —that is, **has the form** $x = a^n b^n$.

   - Well, if $x \in \mathcal{I}$ then $x = \lambda = a^0 b^0$. Done.

   - The property propagates with rule R.

     For example, say X has the property, that is, $X = a^n b^n \in S$. Using (3) we see that the output, aXb, is $a^{n+1} b^{n+1} \in S$. *The property does propagate!* Done.

2. For $S \subseteq \text{Cl}(\mathcal{I}, \mathcal{O})$ we will do induction over $\mathbb{N}$ on the n that occurs in $x = a^n b^n$ (arbitrary member of S) to prove that any $x \in S$ satisfies $\underbrace{x \in \text{Cl}(\mathcal{I}, \mathcal{O})}_{P[x]}$ ("the property $P[x]$").

   We do SI.

   Basis. $n = 0$. Let $x = a^0 b^0$, a member of S. This is equal to λ hence is in $\text{Cl}(\mathcal{I}, \mathcal{O})$ too (in $\mathcal{I}$ in fact).

   I.H. **Assume for fixed** n that $a^n b^n$ of S is in the closure.

   I.S. **Prove now for the same** n**, that** $a^{n+1} b^{n+1}$ **in** S is in the closure as well.

   Well, Our ONLY operation transforms

$$a^n b^n \overset{I.H.}{\in} \text{Cl}(\mathcal{I}, \mathcal{O})$$

   into $aa^n b^n b = a^{n+1} b^{n+1}$. Thus, $a^{n+1} b^{n+1} \in \text{Cl}(\mathcal{I}, \mathcal{O})$ by the closure of this set under $X \mapsto aXb$. Done. □

7.1.6 Example. Regarding the example of simple arithmetic expressions (p.333) we prove that each such expression has equal numbers of left and right brackets.

Here $\mathcal{I} = \{1, 2, 3\}$ and the two operations (on strings) are

$$\begin{matrix} E \\ \longrightarrow \\ \longrightarrow \\ E' \end{matrix} \boxed{\ +\ } \longrightarrow (E + E') \tag{1}$$

and

$$\begin{matrix} E \\ \longrightarrow \\ \longrightarrow \\ E' \end{matrix} \boxed{\ \times\ } \longrightarrow (E \times E') \tag{2}$$

Well, each of the members of $\mathcal{I}$ has the claimed property —0 number of left and 0 number of right brackets.

The property *is preserved by each of (1) and (2)*. For example,

If E and E' have the property (by I.H.) so does "$(E + E')$"

since we added just <u>one</u> left bracket and <u>one</u> right bracket to the already existing brackets of E and E' where —in each of these two strings— the # of left ones and the # of right ones balance out *by the I.H.*

Similarly for $\times$*.* □

Bibliography

[Dav65] M. Davis, *The undecidable*, Raven Press, Hewlett, NY, 1965.

[DS90] Edsger W. Dijkstra and Carel S. Scholten, *Predicate Calculus and Program Semantics*, Springer-Verlag, New York, 1990.

[GS94] David Gries and Fred B. Schneider, *A Logical Approach to Discrete Math*, Springer-Verlag, New York, 1994.

[Kle43] S.C. Kleene, *Recursive predicates and quantifiers*, Transactions of the Amer. Math. Soc. **53** (1943), 41–73, [Also in [Dav65], 255–287].

[Kur63] A.G. Kurosh, *Lectures on General Algebra*, Chelsea Publishing Company, New York, 1963.

[Tou03a] G. Tourlakis, *Lectures in Logic and Set Theory, Volume 1: Mathematical Logic*, Cambridge University Press, Cambridge, 2003.

[Tou03b] ———, *Lectures in Logic and Set Theory, Volume 2: Set Theory*, Cambridge University Press, Cambridge, 2003.

[Tou08] ———, *Mathematical Logic*, John Wiley & Sons, Hoboken, NJ, 2008.