

Set Theory and Foundations of Mathematics

2. Set theory

Sylvain Poirier
<http://settheory.net/>

2.1. First axioms of set theory

The inclusion predicate

The binary predicate \subset of [inclusion](#) between sets is defined by : for all sets E and F ,

$$E \subset F \Leftrightarrow \forall x \in E, x \in F$$

and read as " E is included in F ", or " E is a subset of F ", or " F includes E ". Properties of inclusion between classes apply.

$E \subset E$ is logically valid. [Implications chains](#) also appear as inclusion chains:

$$(E \subset F \subset G) \Leftrightarrow (E \subset F \wedge F \subset G) \Rightarrow E \subset G.$$

We may use the same notation $E \subset F$ as abbreviation for the inclusion of a set E in a class F (and similarly for other formulas defined from this):

$$E \subset F \Leftrightarrow \forall x \in E, F(x)$$

Formulas vs statements

Most set theories (except mainly [FST](#) and some strong versions) will only accept [bounded formulas](#) as sub-formulas of terms (by the set-builder, and later by the [conditional operator](#)) and as possible definitions of predicates (what predicate symbols abbreviate). Open quantifiers will be only allowed in statements (declared true as axioms or theorems).

In set theory, a *statement* is a [ground](#) formula which can combine the symbols of first-order logic with the regular ones of set theory as follows: it must be made of a chain of open quantifiers, usually all \forall and often written in words ("for all"), followed by a bounded formula. Proofs will naturally use the deduction rules for open quantifiers (introduction and elimination) by common language articulations.

The above definition of \subset , basically claiming the predicate symbol \subset to abbreviate the bounded formula $(\forall x \in E, x \in F)$, can also be seen as an example of statement with open quantifiers $(\forall_{\text{Set}} E, F,)$ taken as an axiom.

The role of axioms

As [explained in 1.9](#), the role of the axioms of a generic theory, is to express its intended range of models as a selection ([class](#)) from a wider notion of "all systems" with structures named by the same language. The logical framework, which holds (describes) this wider [notion](#), can interpret the axioms in each such system, then exclude those where it finds some axiom to be false. Excluded systems remain possible models of different theories described by the same framework. Any statement true in all systems (a logically valid statement) is useless (redundant) as an axiom.

But expressing set theory in its special logical framework not used for other theories,

leaves a priori undefined the distinction between its logically valid statements, and its other basic accepted truths which need to be declared as axioms due to their falsity in some "non-universe". Now this distinction can be given sense by converting set theory into a generic theory: let us call *axioms of a set theory* a list of its basic "true" statements whose translated versions in first-order logic are not logically valid, but form a proper list of axioms for that generic theory to be equivalent to the intended version of set theory. From there, *provability* in a set theory can be defined as the one given by first-order logic with these axioms.

Converting the binders

When converting set theoretical expressions into first-order logic, the only modified [symbols](#) are the binders, as their format of use differs between both frameworks. Let us describe the rules of this conversion.

The [function definer](#) (1.8) becomes an infinity of operator symbols: for each term t with one argument and any list of parameters, the whole term $(E \ni x \mapsto t(x))$ is seen as the big name of a distinct operator symbol, whose arguments are E and the parameters of t . (Those where every subexpression of t without any occurrence of x is the only occurrence of a parameter, would suffice to define others).

The same goes for the set-builder, which will [come as a particular case in 2.2](#).

The conversion of quantifiers comes by [expressing their domains as classes](#) :

$$(\exists x \in E, A(x)) \rightarrow (\exists x, x \in E \wedge A(x))$$

$$(\forall x \in E, A(x)) \rightarrow (\forall x, x \in E \Rightarrow A(x))$$

Classification of axioms

Axioms of set theories (sometimes with other primitive components) can be classified as follows according to their roles, ordered from the more "primitive" (necessary) components, to the more optional and debatable ones (opening a diversity of acceptable set theories).

- The *technical axioms* give to the primitive notions and symbols their correct roles :
 - The notions of sets and functions, [symbolized in 1.7](#), are axiomatized here in 2.1 : axioms for notions ; axiom of extensionality ; axioms for functions.
 - Unique element (2.4) ;
- Axioms from the set generation principle (2.2) ;
- Strengthening axioms, [introduced in 1.A](#) ;
- More optional technical axioms will come later:
 - Axiom of choice (2.10) might be seen as a further specification for the powerset ;
 - Axiom of foundation (evoked in 2.A and expressed in 5.3).

Axioms for notions

The formalization of primitive notions by class symbols following [1.7](#), needs to be completed by the following axioms.

$\forall x, \neg(\text{Set}(x) \wedge \text{Fnc}(x))$: sets are not functions (though it does not matter)

$\forall_{\text{Fnc}} f, \text{Set}(\text{Dom } f)$: the domain of every function is a set

$(\text{Fc}) \forall(t, E), \text{Fnc}(E \ni x \mapsto t(x))$: any definite $(E \ni x \mapsto t(x))$ is a function

Here and in the below axioms for functions, $\forall(t, E)$ is meant as declaring an axiom schema by second-order universal elimination over the variable functor t with E included in the definiteness class of the term defining t for the given values of parameters (this is the definiteness condition for $(E \ni x \mapsto t(x))$). Thus for each term defining t we have an axiom

where $\forall(t,E)$ is replaced by

$$\forall_{\text{parameters}}, \forall_{\text{Set}} E, (\forall x \in E, dt(x, \text{parameters})) \Rightarrow$$

Axiom of Extensionality

This axiom lets sets be determined by their role (either as ranges for quantifiers, or as the classes they define by \in), saying that any two sets playing the same role are equal :

$$\forall_{\text{set}} E, F, E \subset F \subset E \Rightarrow E = F.$$

Indeed, $E \subset F \subset E$ means that E and F have the same elements ($\forall x, x \in E \Leftrightarrow x \in F$), and for any predicate R ,

$$(\forall x \in F, R(x)) \Leftrightarrow (\forall x \in E, R(x))$$

and similarly for \exists . Informally, the elements of a set are given in bulk.

Axioms for functions

Functions are also made determined by their role, by the following axiom

$$(\text{=}_{\text{Fnc}}) : \forall_{\text{Fnc}} f, g, (\text{Dom } f = \text{Dom } g \wedge \forall x \in \text{Dom } f, f(x) = g(x)) \Rightarrow f = g.$$

The function definers are related with the function evaluator by an axiom which can be written in either way:

1. $\forall(t,E), \forall_{\text{Fnc}} f, f = (E \ni x \mapsto t(x)) \Leftrightarrow (\text{Dom } f = E \wedge \forall x \in E, f(x) = t(x))$
2. $\forall(t,E), \text{Dom}(E \ni x \mapsto t(x)) = E \wedge \forall x \in E, (E \ni y \mapsto t(y))(x) = t(x)$

Indeed $(\text{Fc}) \Leftrightarrow (1. \Leftrightarrow (2. \wedge (\text{=}_{\text{Fnc}})))$.

► The proof can be taken as an exercise (click to show)

Assume (Fc) for 2. to be definite.

Abbreviate $(\text{Dom } f = E \wedge \forall x \in E, f(x) = t(x))$ as $[f : E, t]$, to shorten the above as

1. $\forall(t,E), \forall_{\text{Fnc}} f, f = (E \ni x \mapsto t(x)) \Leftrightarrow [f : E, t]$
2. $\forall(t,E), [E \ni x \mapsto t(x) : E, t]$

1. $\Leftrightarrow (2. \wedge 1b.)$ where

$$1b. \forall(t,E), \forall_{\text{Fnc}} f, [f : E, t] \Rightarrow f = (E \ni x \mapsto t(x))$$

Proof of $1b. \Rightarrow (\text{=}_{\text{Fnc}})$

$$\begin{aligned} & \forall_{\text{Fnc}} g, \text{ applying } 1b. \text{ to } (g, E) \text{ where } E = \text{Dom } g \text{ as } \forall x \in \text{Dom } g, dg(x) : \\ & \forall_{\text{Fnc}} f, [f : E, g] \Rightarrow f = (E \ni x \mapsto g(x)) \\ & [g : E, g] \therefore g = (E \ni x \mapsto g(x)) \\ & \forall_{\text{Fnc}} f, [f : E, g] \Rightarrow f = g \end{aligned}$$

Proof of $(2. \wedge (\text{=}_{\text{Fnc}})) \Rightarrow 1b.$

$$\forall(t,E), \forall_{\text{Fnc}} f, ([f : E, t] \wedge [E \ni x \mapsto t(x) : E, t]) \Rightarrow f = (E \ni x \mapsto t(x)). \blacksquare$$

Note: $(\text{Fc}) \wedge 1.$ can also be written

$$1'. \forall(t,E), \forall f, f = (E \ni x \mapsto t(x)) \Leftrightarrow (\text{Fnc } f \wedge \text{Dom } f = E \wedge \forall x \in E, f(x) = t(x)).$$

2.2. Set generation principle

Formalizing diverse notions in set theory

Set theory will be extended by additional components to represent diverse kinds of meta-objects (roles played by symbolic systems) directly as objects (which can play them). This can only apply to meta-objects which are "concrete" i.e. object-like enough, in a sense to be clarified : as seen in 1.8 and 1.D, not all classes can be sets, while only the functors whose domains are sets are represented by functions.

Some of those extensions will consist in primitive "definer" symbols (with the relevant axioms) serving to convert into an existing kind of objects some meta-objects (roles) of the same kind :

- Diverse "set-like" classes will be converted into sets (2.2) (thus giving the existence of sets coinciding with these classes);
- Indirectly specified elements will become directly designated (2.4).

Other kinds of meta-objects beyond the normal roles of elements, sets and functions, will form new notions : those of operation, relation and tuple will be formalized in 2.3. The new symbols they need are both definers (converting meta-objects into the new objects) and evaluators (interpreting these objects back to their roles as meta-objects).

But, like for unary relations (1.8), these new notions with their role-giving symbols, need not be primitive, as they can be naturally *constructed* as (represented by) classes of existing objects (sets or functions), which can offer their tools (expressible features) to the new objects. So, these extensions of set theory can be obtained by giving definitions of their definer and evaluator symbols. As the notions of operation and relation can be represented in several useful ways by classes of old objects, which do not represent each new object by the same old object (2.3), functors will be used for translations between the classes playing the roles of these notions in the different ways (2.6, 2.8).

Sets as domains of bounded quantifiers

Bounded quantifiers give sets their fundamental role as ranges of bound variables, unknown by the predicate \in which only gives them their role of classes. Technically, the bounded quantifier $(\exists \in ,)$ suffices to define the predicate \in as

$$x \in E \Leftrightarrow (\exists y \in E, x = y)$$

but is not generally definable from it in return by any bounded formula, as [its definition](#) involves an open quantifier.

Philosophically speaking, the perception of a set as a class (ability to classify each object as belonging to it or not) does not suffice to provide its full perception as a set, that is the perspective over all its elements as coexisting. This intuitive view of a set as a coexistence of its elements can be seen as supported by the above definition of membership as a form of existence with respect to this set. This way, Russell's paradox means that mathematical objects cannot all coexist as a ultimate totality.

Statement of the principle

For any data of a class C and a quantifier Q [defined by](#) given bounded formulas with the same parameters, if the equivalence between Q and \forall_C , namely

$$\forall A, (\forall x, C(x) \Rightarrow A(x)) \Leftrightarrow (Qx, A(x))$$

is proven by [second-order universal introduction](#) (*) under a given condition on parameters, then C is a set that can be named by a new operator symbol K added to the language of set theory, with these parameters as arguments, and the following axiom

which expresses $K = C$ by [double inclusion](#): for all values of parameters satisfying the condition,

$$\text{Set}(K) \wedge (\forall x \in K, C(x)) \wedge (Qx, x \in K).$$

(*) we shall focus on cases whose proofs do not use any axiom obtained from a second order axiom by using A in the definition of its variable structure, though I am not sure if any such use is possible.

This principle will be justified in [2.B](#).

The equivalence condition between Q and \forall_C is equivalently expressible as the following list of 3 statements, where the quantifier Q^* defined by $(Q^*x, A(x)) \Leftrightarrow (Qx, \neg A(x))$ will be equivalent to \exists_C :

- (1) $\forall x, C(x) \Leftrightarrow Q^*y, x=y$ (in fact we just need $\forall x, C(x) \Rightarrow Q^*y, x=y$)
- (2) $Qx, C(x)$
- (3) $\forall A, B, (\forall x, A(x) \Rightarrow B(x)) \Rightarrow ((Qx, A(x)) \Rightarrow (Qx, B(x)))$.

► Proof of equivalence (click to show)

These 3 properties are already known consequences of « $Q = \forall_C$ ». Conversely,

$$((2) \wedge (3)) \Rightarrow ((\forall_C x, A(x)) \Rightarrow Qx, A(x))$$

$$((1) \wedge \exists_C x, A(x)) \Rightarrow \exists y, (Q^*x, x=y) \wedge (\forall x, x=y \Rightarrow A(x)) \therefore ((3) \Rightarrow Q^*x, A(x)) \blacksquare$$

Main examples

Our first and main use cases of this principle are listed in the following table. To justify all of these, (1) comes by finding C to be the formula so defined from Q , while (3) is ensured by the absence of any "negative" occurrence of A in Q (inside a negation, an equivalence, or left of a \Rightarrow). This leaves (2) as the easy remaining condition to check.

The first column recalls the above generic abbreviations ; each other column gives an effective example. The K line gives the notations for the set theoretical symbols so introduced.

K	$\{y \in E \mid B(y)\}$	$\bigcup E$	$\text{Im } f$	\emptyset	$\{y\}$	$\{y, z\}$
$C(x)$	$x \in E \wedge B(x)$	$\exists y \in E, x \in y$	$\exists y \in \text{Dom } f, f(y) = x$	0	$x = y$	$x = y \vee x = z$
dK	$\forall x \in E, dB(x)$	$\text{Set}(E) \wedge \forall y \in E, \text{Set}(y)$	$\text{Fnc}(f)$	1	1	1
$Qx, A(x)$	$\forall x \in E, B(x) \Rightarrow A(x)$	$\forall y \in E, \forall x \in y, A(x)$	$\forall x \in \text{Dom } f, A(f(x))$	1	$A(y)$	$A(y) \wedge A(z)$
$Q^*x, A(x)$	$\exists x \in E, B(x) \wedge A(x)$	$\exists y \in E, \exists x \in y, A(x)$	$\exists x \in \text{Dom } f, A(f(x))$	0	$A(y)$	$A(y) \vee A(z)$

The value of the set-builder $K = \{x \in E \mid B(x)\}$ first defined in 1.8 as a class, thus with an open quantifier ($\forall x, x \in K \Leftrightarrow (x \in E \wedge B(x))$) is now re-defined by axioms without open quantifier beyond parameters:

$$\text{Set}(K) \wedge (\forall x \in K, x \in E \wedge B(x)) \wedge (\forall x \in E, B(x) \Rightarrow x \in K)$$

or more shortly $\text{Set}(K) \wedge K \subset E \wedge (\forall x \in E, x \in K \Leftrightarrow B(x))$. Like the axioms for functions, this is an axiom schema by second-order universal elimination on $\forall B$, so with all unary predicates B defined by bounded formulas with parameters.

This way, the proof of Russell's paradox would be written

$$\exists F, (F = \{x \in E \mid \text{Set}(x) \wedge x \notin x\}) \therefore \text{Set}(F) \therefore (F \in F \Leftrightarrow (\text{Set}(F) \wedge F \notin F)) \therefore F \notin E$$

The functor \bigcup is the union symbol, and its axioms form the *axiom of union*.

The set $\text{Im } f$ of values of $f(x)$ when x ranges over $\text{Dom } f$, is called the *image* of f .

We define the predicate $(f: E \rightarrow F)$ that reads « f is a function from E to F », while E and F are sets (definiteness condition which may also be integrated in this definition), as

$$(f: E \rightarrow F) \Leftrightarrow (\text{Fnc}(f) \wedge \text{Dom } f = E \wedge \text{Im } f \subset F)$$

A set F such that $\text{Im } f \subset F$ (i.e. $\forall x \in \text{Dom } f, f(x) \in F$), is called a *target set* of f . The more precise $(\text{Fnc}(f) \wedge \text{Dom } f = E \wedge \text{Im } f = F)$ will be denoted $f: E \rightarrow F$ (f is a *surjection*, or *surjective function* from E to F , or function from E onto F).

The *empty set* \emptyset is the only set without element, and is included in any set ($\forall_{\text{Set } E}, \emptyset \subset E$). Thus, $(E = \emptyset \Leftrightarrow E \subset \emptyset \Leftrightarrow \forall x \in E, 0)$, and thus $(E \neq \emptyset \Leftrightarrow \exists x \in E, 1)$.

The presence of this constant symbol \emptyset ensures the existence of a set; for any set E we also get $\emptyset = \{x \in E \mid 0\}$. We have $\bigcup \emptyset = \emptyset$.

As $(\text{Dom } f = \emptyset \Leftrightarrow \text{Im } f = \emptyset)$ and $(\text{Dom } f = \text{Dom } g = \emptyset \Rightarrow f = g)$, the only function with domain \emptyset is called the *empty function* \emptyset_{\mapsto} .

That a variable may have empty range, is no obstacle for fixing it. In particular, developing an inconsistent theory means studying a fixed system whose range of possibilities is empty. We may actually need to do so in order to discover a contradiction there, which means to prove that no such system exists.

We can redefine \exists from the above in two ways:

$$(\exists x \in E, A(x)) \Leftrightarrow \{x \in E \mid A(x)\} \neq \emptyset \Leftrightarrow (1 \in \text{Im}(E \ni x \mapsto A(x))).$$

For all x , $\{x, x\} = \{x\}$. Such a set with only one element is called a *singleton*.

For all x and y we have $\{x, y\} = \{y, x\}$. If $x \neq y$, the set $\{x, y\}$ with 2 elements x and y is called a *pair*.

For any set E , the *identity function* on E is defined by

$$\text{Id}_E = (E \ni x \mapsto x).$$

Thus,

$$\begin{aligned} \text{Dom } \text{Id}_E &= E = \text{Im } \text{Id}_E \\ \forall x \in E, \text{Id}_E(x) &= x. \end{aligned}$$

2.3. Currying and tuples

Formalizing operations and relations

To formalize in set theory every notion of [operation](#) (resp. [relation](#)) with any fixed arity $n > 1$, acting as n -ary structures between n sets, would involve the following symbols (generalizing the case $n = 1$ of functions which we took as primitive):

- A list of n domain functors (sometimes not available but not needed since domains are usually fixed by the context);
- Its evaluator is an $(n+1)$ -ary operator (resp. predicate) written in filled form as $f(x_0, \dots, x_{n-1})$, with arguments one n -ary operation (resp. relation) f , and its n arguments x_0, \dots, x_{n-1} ;
- Its definer binds n variables to their respective ranges on a term (resp. a formula).

Let us give to the operation definers, the following notation which will be later (2.7) justified as reflecting their definitions: in the binary case ($n = 2$), operations defined by a binary operator A with domains E and F will be written

$$E \times F \ni (x, y) \mapsto A(x, y)$$

After seeing what operations and relations are (which role they play) let us discuss how they are reducible to something else.

One aspect is that relations are reducible to operations, and vice-versa:

- The notion of n -ary relation can be constructed as a class of n -ary operations, by taking the Boolean pair as target set, as we did in [1.4](#) (for predicates) and will formalize in 2.4.
- Inversely, the notion of n -ary operation can be constructed as a class of $(n+1)$ -ary relations (2.7).

These constructions can be useful despite their circularity (they are not decreasing the arity). But for foundational purposes, we need a construction which reduces to lower arity cases, down to sets and functions. There are actually 2 such constructions.

Currying

The simplest constructions of the notions of n -ary operation from our primitive tools for functions, are their curried representations, defined as classes of functions by expressing the operation definer as n successive uses of the function definer (one for each bound variable); and thus the evaluator as n uses of the function evaluator. For $n=2$, this means a binary operation in curried form b with domains E and F , is defined from a binary operator B , and evaluated back as B , by the formulas

$$b = (E \ni x \mapsto (F \ni y \mapsto B(x,y)))$$

$$\text{Dom } b = E \wedge \forall x \in E, \text{Dom } b(x) = F \wedge \forall y \in F, B(x,y) = b(x)(y)$$

whose equivalence, similar to the [axiom 1. for functions](#), is deducible from it.

The first domain functor comes as $\text{Dom } b = E$, but the second one, given as common value of all $\text{Dom } b(x)$, is only definite for $E \neq \emptyset$. The intermediate function $b(x) = (F \ni y \mapsto B(x,y))$ represents the functor defined by B with parameter x (free) and argument y (bound).

This somehow makes each binary operation work as a function evaluator, by giving x the role of a function with argument y ; the curried function evaluator just gives back each function it takes : $f \mapsto ((\text{Dom } f) \ni y \mapsto f(y))$.

Similarly, the notion of n -ary relation can be defined in curried form as a class, giving the role of definer to a succession of $n-1$ uses of the function definer and 1 use of the set-builder. The curried form r of a binary relation with domains E and F , is so defined from a binary predicate R , and evaluated back as R , by the equivalent formulas

$$r = (E \ni x \mapsto \{y \in F \mid R(x,y)\})$$

$$\text{Dom } r = E \wedge (\forall x \in E, r(x) \subset F \wedge \forall y \in F, R(x,y) \Leftrightarrow y \in r(x))$$

This leaves no way to define second domain functor (F cannot be restored from r).

These constructions have the defect of breaking the symmetry of role between arguments. The next construction method, by tuples, will preserve this symmetry, and will be usually preferred for this reason.

Tuples

For any theory and any natural number n , a notion of n -tuple is a notion constructed for its variables to serve as abbreviations for a pack of n variables from previous notions. A 2-tuple is called an *ordered pair*, a 3-tuple is a *triple*, a 4-tuple is a *quadruple* and so on. So each tuple x (so constructed object) represents a meta-function of "interpretation" from this list (meta-set) of n variables, into the model.

The definer and evaluator symbols for tuples take a special form because, while tuples are meta-functions, the role they play in the theory differs from those of functors or functions.

An n -tuple definer is not a binder but an n -ary operator, packing its n variables into one by taking them as its arguments in a parenthesis and separated by commas: (...,).

Now, each structure (resp. operation, relation) can be re-formalized as a unary structure (resp. a function, a unary relation) with domain a type (resp. a class which is actually a

set) of tuples, to be used with the relevant tuples definer. This can be implicitly done by re-interpreting the notation: $S(x,y)$ can be read seeing S either as a binary structure (resp. operation, relation), or as a unary one taking as argument the ordered pair given by the term (x,y) formed by the ordered pair definer with arguments x and y .

The role of n -tuple evaluator is played by a list of n functors called *projections* which extract each variable from the given pack. This list plays the role of the function evaluator carried in the reverse way, by fixing its meta-argument whose range (the domain of tuples as meta-functions) is not seen by the theory as a set of objects but as a list of n symbols to be taken separately.

The formal details of this construction depend whether we take a generic theory or a set theory.

Tuples axiom

In generic theories, it must be a distinct constructed type for not only each n but also each choice of the type of each of these n variables (if starting from a theory with multiple types - this choice of a list of variables with their types amounts to the choice of a [predicate symbol type](#)), but this directly forms a [legitimate construction](#) (4.11).

For $n=2$ (other cases are similar), and a type P of ordered pairs (x,y) where x,y have types E and F , the projections π_0 and π_1 are related to the ordered pair definer, by the axiom

$$\forall_E x, \forall_F y, \forall_P z, (x = \pi_0(z) \wedge y = \pi_1(z)) \Leftrightarrow (x,y) = z$$

which implies

$$\begin{aligned} \forall_P z, (\pi_0(z), \pi_1(z)) &= z \\ \forall_E x, \forall_F y, \pi_0(x,y) &= x \\ \forall_E x, \forall_F y, \pi_1(x,y) &= y \\ \forall_E x,x', \forall_F y,y', (x,y) &= (x',y') \Leftrightarrow (x = x' \wedge y = y') \end{aligned}$$

Tuples in set theory

Set theory having no types only needs one notion of n -tuple for each n . But it faces one difficulty : in order to rigorously justify that all tools of set theory remain applicable to tuples in the same way as to the rest of elements, tuples need to be integrated among elements. Namely, set theory needs to develop its notions of tuples by **defining** them as classes of already existing elements, instead of simply accepting them by the above **constructions**, which would leave them as separate types.

For this, since tuples are meta-functions, they are naturally represented by functions. This involves copying their meta-domain, which is a list of symbols, into a set (of elements). Namely, let us represent the domain of considered n -tuples by the set V_n of n numbers from 0 to $n-1$ all named by constant symbols we shall simply conceive as digits. To justify this as a development of set theory, these digits can be defined by any ground terms with distinct values, such as \emptyset , $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$ and endlessly many more sets and functions expressed from there. The notion of n -tuple is then defined as the class of all functions with domain V_n .

Then, the projections are defined by the function evaluator applied to the respective digits, which can be read as disguised in their notations : for all n -tuples x ,

$$\forall i \in V_n, \pi_i(x) = x(i).$$

This index notation, making a pun of the equivalence between reading the index i as an argument or a meta-argument that means to take the i -th symbol from a list (here π_i is the

i -th projection), is also commonly used for tuples : x_i can be either read as the i -th variable in the tuple $x = (x_0, \dots, x_{n-1})$, or as a third notation for the same evaluator after $x(i)$ and $\pi_i(x)$.

More tools need to be introduced before defining in set theory the tuple definers (2.4) and the operation definers working with tuples (2.7).

2.4. Uniqueness quantifiers and conditional operator

Uniqueness quantifiers

For all sets $F \subset E$, all unary predicate A definite on E , and all $x \in E$,

$$\begin{aligned} x \in F &\Leftrightarrow \{x\} \subset F \Leftrightarrow (\exists y \in E, x=y \wedge y \in F) \Leftrightarrow (\forall y \in E, x=y \Rightarrow y \in F) \\ x \in F &\Rightarrow ((\forall y \in F, A(y)) \Rightarrow A(x) \Rightarrow \exists y \in F, A(y)) \\ F \subset \{x\} &\Leftrightarrow (\forall y \in F, x=y) \Leftrightarrow ((\exists y \in F, A(y)) \Rightarrow A(x) \Rightarrow (\forall y \in F, A(y))) \\ F = \{x\} &\Leftrightarrow (x \in F \wedge \forall y \in F, x=y) \Leftrightarrow (\forall y \in E, y \in F \Leftrightarrow x=y) \end{aligned}$$

Here are 3 new quantifiers: $\exists^{\geq 2}$ (plurality), $!$ (uniqueness), and $\exists!$ (existence and uniqueness), whose results when applied to A in E only depend on $F = \{x \in E \mid A(x)\}$ (like \exists and unlike \forall) :

$$\begin{aligned} (\exists x \in E, A(x)) &\Leftrightarrow (F \neq \emptyset) \Leftrightarrow (\exists x \in F, 1) \Leftrightarrow (\exists x \in E, \{x\} \subset F) \\ (\exists^{\geq 2} x \in E, A(x)) &\Leftrightarrow (\exists^{\geq 2}: F) \Leftrightarrow (\exists x, y \in F, x \neq y) \Leftrightarrow (\exists x \neq y \in E, A(x) \wedge A(y)) \\ (!x \in E, A(x)) &\Leftrightarrow (!: F) \Leftrightarrow \neg(\exists^{\geq 2}: F) \Leftrightarrow (\forall x, y \in F, x=y) \Leftrightarrow \forall x \in F, F \subset \{x\} \\ (\exists! x \in E, A(x)) &\Leftrightarrow (\exists!: F) \Leftrightarrow (\exists x \in F, F \subset \{x\}) \Leftrightarrow (\exists x \in E, F = \{x\}) \\ F \subset \{x\} &\Leftrightarrow (\forall y \in F, F \subset \{y\}) \Leftrightarrow (!: F) \\ (\exists!: F) &\Leftrightarrow (F \neq \emptyset \wedge !: F) \\ F \neq \emptyset &\Leftrightarrow ((\forall y \in F, B(y)) \Rightarrow (\exists y \in F, B(y))) \\ (!: F) &\Leftrightarrow ((\exists y \in F, B(y)) \Rightarrow (\forall y \in F, B(y))) \\ F = \{x\} &\Leftrightarrow ((\exists y \in F, B(y)) \Rightarrow B(x) \Rightarrow \forall y \in F, B(y)) \end{aligned}$$

Single element axiom

Let us introduce a new primitive symbol ι to serve as inverse of the [singleton functor](#) $\{ \}$: only definite on the class of singletons $\text{d}_1 E \Leftrightarrow (\text{Set}(E) \wedge \exists!: E)$, it gives their unique element, as defined by the *single element axiom* which can be equivalently written

$$\begin{aligned} \forall x, \iota\{x\} &= x \\ \forall_{\text{Set} E}, \exists!: E &\Rightarrow \iota E \in E \end{aligned}$$

Then for every unary predicate A and every singleton E , $A(\iota E) \Leftrightarrow (\exists x \in E, A(x)) \Leftrightarrow (\forall x \in E, A(x))$.

Conditional connective

Following conventions from computer science, let us denote the ternary connective «*If A then B otherwise C*» as

$$\begin{aligned} (A ? B : C) &\Leftrightarrow (\neg C \Rightarrow A \Rightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (\neg A \Rightarrow C)) \Leftrightarrow (\neg A ? C : B) \\ &\Leftrightarrow ((A \wedge B) \vee (\neg A \wedge C)) \Leftrightarrow ((C \Rightarrow A) \Rightarrow (A \wedge B)) \Leftrightarrow (A ? \neg B : \neg C) \end{aligned}$$

Any $n+1$ -ary connective K amounts to two n -ary ones: $K(A) \Leftrightarrow (A ? K(1) : K(0))$. Thus,

$$\begin{aligned}
\neg A &\Leftrightarrow (A ? 0 : 1) \\
(A \Rightarrow B) &\Leftrightarrow (A ? B : 1) \\
(A \wedge B) &\Leftrightarrow (A ? B : 0) \\
(A \vee B) &\Leftrightarrow (A ? 1 : B) \\
(A \Leftrightarrow B) &\Leftrightarrow (A ? B : \neg B).
\end{aligned}$$

Conditional operator

Like the conditional connective, it chooses between two objects x,y depending on the boolean B :

$$(B ? x : y) = \iota\{z \in \{x,y\} \mid B ? z=x : z=y\} = \iota\{z \in \{x,y\} \mid z=x \Leftrightarrow B\}$$

so that for any unary predicate A ,

$$A(B ? x : y) \Leftrightarrow (B ? A(x) : A(y)).$$

All para-operators other than connectives but with at least a Boolean argument, are naturally expressed as composed of the conditional operator with operators, or the conditional connective with predicates, which is why they did not need an explicit presence in the language of a theory.

Relations as operations

For a given n , once given the notion of n -ary operation with its definer and its evaluator denoted as in 2.3, the notion of n -ary relation can be defined as a role of the class of n -ary operations with values in $V_2 = \{0,1\}$ (serving as Boolean type), as follows. Booleans B are converted to and from objects $b \in V_2$, by $b = (B ? 1 : 0)$ and $B \Leftrightarrow b = 1$.

So with $n=2$ (other cases are similar), such a binary operation r with domains E and F gets this role of binary relation, being defined from any binary predicate R , and evaluated back as R , by the formulas

$$\begin{aligned}
r &= (E \times F \ni (x,y) \mapsto (R(x,y) ? 1 : 0)) \\
\forall x \in E, \forall y \in F, R(x,y) &\Leftrightarrow r(x,y) = 1
\end{aligned}$$

Defining the tuples definers

The ordered pair definer can be defined as a curried view on the conditional operator:

$$(x,y) = (V_2 \ni a \mapsto (a = 1 ? y : x))$$

Confusing Booleans with objects, $(B ? x : y) = \pi_B(y,x)$.

Other tuples can be similarly defined in diverse ways, such as

$$(x,y,z) = (V_3 \ni a \mapsto (a = 2 ? z : (a = 1 ? y : x)))$$

Given the notion of ordered pairs, the roles of notions of tuples of any higher arity $n > 2$ can also be played by other classes, with other definitions of their definers and projections, satisfying the same axioms. For example, triples $t = (x,y,z)$ can be defined as $t = ((x,y),z)$ and evaluated by

$$\begin{aligned}
x &= \pi_0(\pi_0(t)) \\
y &= \pi_1(\pi_0(t)) \\
z &= \pi_1(t)
\end{aligned}$$

2.5. Families and Boolean operators on sets

Families

The notion of *family* is formally synonymous with that of function, but meant to play roles intuitively similar to those of tuples : while the domains of tuples are finite and made of symbols or numbers, those of families can be more general but still seen as intuitively "simple" sets, fixed in each context of use, and outside the "main studied system" containing the target set. A «family of...» is a family whose image is a «set of... » (included in the said notion).

Families use the formalism of functions disguised in the style of the symbols for tuples (themselves inapplicable due to their finiteness for infinite domains). The evaluator for a family u at i , instead of $u(i)$, is written $\pi_i(u)$ or u_i (looking like a meta-variable symbol of variable). The family definer on a term t is written $(t(i))_{i \in I}$ or sometimes with ellipsis as $(t(0), \dots, t(n-1))$ abbreviating the n -tuple definer, instead of $I \ni i \mapsto t(i)$. The argument i is called *index*, and the family is said *indexed* by its domain I . We may write $u = (u_i)_{i \in I}$ as a shorthand for $\text{Fnc } u \wedge \text{Dom } u = I$.

When formalizing model theory in set theory, the "lists" of components forming the [contents of theories](#) are basically families. More precisely:

- The list of abstract types is a set of urelements, serving as indexing set (i.e. domain) for the family of interpreted types
- Both lists of predicate symbols and operator symbols, are families of tuples of abstract types
- Of the list of axioms, only its image matters, which is a subset of the set of all statements.

Structures and binders

Seeing tuples as particular functions, explains that structures ([unary structures on classes of tuples](#)) are particular binders over terms (unary structures on classes of functions defined by these terms: [1.8](#)).

Similarly, [logical connectives](#) are particular [quantifiers](#). In particular, \forall and \exists respectively generalize the chains of conjunctions and of disjunctions:

$$(B_0 \wedge \dots \wedge B_{n-1}) \Leftrightarrow (\forall i \in V_n, B_i)$$
$$(B_0 \vee \dots \vee B_{n-1}) \Leftrightarrow (\exists i \in V_n, B_i)$$

The distributivity of connectives (1.6) also works for quantifiers as follows. For any unary predicate R definite on E and any Boolean C ,

$$(C \wedge \exists x \in E, R(x)) \Leftrightarrow (\exists x \in E, C \wedge R(x))$$
$$(C \vee \forall x \in E, R(x)) \Leftrightarrow (\forall x \in E, C \vee R(x))$$
$$(C \Rightarrow \forall x \in E, R(x)) \Leftrightarrow (\forall x \in E, C \Rightarrow R(x))$$
$$((\exists x \in E, R(x)) \Rightarrow C) \Leftrightarrow (\forall x \in E, R(x) \Rightarrow C)$$
$$(\exists x \in E, C) \Leftrightarrow (C \wedge E \neq \emptyset) \Rightarrow C \Rightarrow (C \vee E = \emptyset) \Leftrightarrow (\forall x \in E, C)$$

Extensional definitions of sets

The operators of empty set \emptyset , singleton $\{a\}$ and pairing $\{a,b\}$ ([2.2](#)) are the first cases (those of arities 0,1,2) of operator symbols of *extensional definition* of sets (listing their elements) which amount to applying the Im functor to tuples :

$$\{a,b,\dots\} = \text{Im}(a,b,\dots).$$

The definition of V_n from 2.3 can be written $\{0,\dots,n-1\}$. Images of tuples are finite sets. Similar notations are used for the binder also given by Im :

$$\{T(x) \mid x \in E\} = \{T(x)\}_{x \in E} = \text{Im}(E \ni x \mapsto T(x))$$

As this notation looks similar to the set builder, both may be combined into a third notation :

$$\{T(x) \mid x \in E \wedge R(x)\} = \{T(x) \mid x \in \{y \in E \mid R(y)\}\}$$

A function f is said *constant* when $! : \text{Im } f$. The constancy of a tuple is the chain of equalities:

$$x=y=z \Leftrightarrow ! : \{x,y,z\} \Leftrightarrow ((x=y) \wedge (y=z)) \Rightarrow x=z$$

Union of a family of sets

The binder of *union* of any family of sets ($F = (F_i)_{i \in I} \wedge \forall i \in I, \text{Set } F_i$), is defined from the previous functor of [union of a set of sets](#) (2.2) :

$$\begin{aligned} \bigcup F &= \bigcup_{i \in I} F_i = \bigcup \{F_i\}_{i \in I} = \bigcup \text{Im } F \\ \forall x, x \in \bigcup_{i \in I} F_i &\Leftrightarrow (\exists X \in \text{Im } F, x \in X) \Leftrightarrow \exists i \in I, x \in F_i \\ X \cup Y &= \bigcup \{X, Y\} \\ \forall x, x \in X \cup Y &\Leftrightarrow (x \in X \vee x \in Y) \\ X \cup Y \cup Z &= \bigcup \{X, Y, Z\} = (X \cup Y) \cup Z \\ \forall_{\text{fnc}} f, \text{Im } f &= \bigcup_{x \in \text{Dom } f} \{f(x)\} \\ (\forall x \in \bigcup_{i \in I} F_i, A(x)) &\Leftrightarrow \forall i \in I, \forall x \in F_i, A(x) \\ (\exists x \in \bigcup_{i \in I} F_i, A(x)) &\Leftrightarrow \exists i \in I, \exists x \in F_i, A(x) \\ \forall_{\text{Set}} E, \bigcup_{i \in I} F_i \subset E &\Leftrightarrow \forall i \in I, F_i \subset E \end{aligned}$$

In return, the union of a set of sets is expressible from that of a family : $\bigcup E = \bigcup \text{Id}_E$ because $\text{Im } \text{Id}_E = E$.

Other Boolean operators on sets

A binary predicate with domains a set I and a class C , can be seen in either curried way, as a meta-family $(A_i)_{i \in I}$ of unary predicates definite in C , or as a family of Boolean variables $(A_i(x))_{i \in I}$ depending on a common parameter x with range C . This way, any quantifier Q with range I (thus, for finite I , any connective) defines a meta-binder (resp. a meta-operation) on this meta-family (resp. this meta-tuple) of unary predicates (sub-classes of C), with result a unary predicate R , defined as

$$\forall_C x, R(x) \Leftrightarrow Q_i, A_i(x).$$

When C is a set E , this defines a genuine binder (resp. operation) acting in the class of subsets of E . However, the result depends a priori not only on the given family of sets but also on the chosen set E . This parameter may be kept implicit as fixed in some contexts, but in others it needs to be explicitly added as argument.

The functor so defined by the negation connective (\neg) is the *complement* C_E (with set parameter E) : $C_E F$ is called the *complement of F in E* . It is synonymous with the following binary operator \setminus of *difference* between any two sets, except that $C_E F$ is considered only definite if $F \subset E$:

$$\begin{aligned} \forall_{\text{Set}} E, F, E \setminus F &= \{x \in E \mid x \notin F\} \\ \forall_{\text{Set}} E, F, F \subset E \Rightarrow C_E F &= E \setminus F \subset E \therefore (\forall x \in E, x \in C_E F \Leftrightarrow \neg x \in F) \end{aligned}$$

Let us analyze the issue of the dependence on E in the general case. Any family of sets $F = (F_i)_{i \in I}$ is a family of subsets of its union $U = \bigcup_{i \in I} F_i$ and of any larger set $E \supset U$.

In any of these, the existential quantifier ($Q = \exists$) defines the union binder:

$$(\forall i \in I, F_i \subset E) \Rightarrow \bigcup_{i \in I} F_i = \{x \in E \mid \exists i \in I, x \in F_i\}$$

The binder defined by any Q gives a set $X = \{x \in E \mid R(x)\}$ where $R(x)$ abbreviates $(Q_i, x \in F_i)$. For this to define a binder (resp. operation) on the class of all sets, this X must be independent of the chosen $E \supset U$, and then it can be written

$$X = \{x \in U \mid R(x)\}.$$

Noting that

$$\forall x, x \in U \vee (R(x) \Leftrightarrow Qi, 0)$$

the needed independence condition appears to be $\neg Qi, 0$ which also implies that X coincides with the class R .

The difference operator can be so defined by the connective $(A \wedge \neg B)$ which satisfies this condition $\neg(0 \wedge \neg 0)$:

$$\forall_{\text{Set}} E, F, \forall x, x \in E \setminus F \Leftrightarrow (x \in E \wedge x \notin F).$$

Intersection

The quantifier \forall defines the *intersection* of any family of subsets F_i of a fixed set E :

$$\bigcap_{i \in I} F_i = \{x \in E \mid \forall i \in I, x \in F_i\} = C_E \bigcup_{i \in I} C_E F_i$$

The condition of independence from E is that $I \neq \emptyset$. As above defined, the intersection of the empty family gives E . The intersection of any non-empty family of sets ($I \neq \emptyset$) can be written

$$\begin{aligned} \forall j \in I, \bigcap_{i \in I} F_i &= \{x \in F_j \mid \forall i \in I, x \in F_i\} \\ \forall x, x \in \bigcap_{i \in I} F_i &\Leftrightarrow \forall i \in I, x \in F_i \\ \forall_{\text{Set}} G, G \subset \bigcap_{i \in I} F_i &\Leftrightarrow \forall i \in I, G \subset F_i \\ \forall_{\text{Set}} E, F, E \cap F &= \{x \in E \mid x \in F\} = F \cap E \\ \forall x, x \in E \cap F &\Leftrightarrow (x \in E \wedge x \in F) \\ E \cap F &\subset E \subset E \cup F \\ E = E \cup F &\Leftrightarrow F \subset E \Leftrightarrow F = E \cap F \end{aligned}$$

Union and intersection have the same associativity and distributivity properties as \wedge and \vee :

$$\begin{aligned} E \cap F \cap G &= (E \cap F) \cap G = E \cap (F \cap G) = \bigcap(E, F, G) \\ E \cap \bigcup_{i \in I} F_i &= \bigcup_{i \in I} E \cap F_i & E \cup \bigcap_{i \in I} F_i &= \bigcap_{i \in I} E \cup F_i \\ E \cap (F \cup G) &= (E \cap F) \cup (E \cap G) & E \cup (F \cap G) &= (E \cup F) \cap (E \cup G) \end{aligned}$$

Finally the connective \oplus gives the symmetric difference: $E \Delta F = (E \cup F) \setminus (E \cap F)$.

2.6. Graphs

A *graph* is a set of ordered pairs. Let us denote the class of graphs as gr :

$$\text{gr } R \Leftrightarrow (\text{Set } R \wedge \forall x \in R, \text{Fnc } x \wedge \text{Dom } x = V_2).$$

For any binder Q and any graph G , the formula $Qz \in G, A(z_0, z_1)$ that binds the variable $z = (z_0, z_1)$ on a binary structure A definite on G , can be seen as binding 2 variables z_0, z_1 on $A(z_0, z_1)$, and thus be denoted with a pair of variables: $Q(x, y) \in G, A(x, y)$.

The *transpose* of an ordered pair is

$${}^t(x, y) = (y, x)$$

The transpose of a graph R is the image of transposition over it:

$${}^tR = \{(y, x) \mid (x, y) \in R\}$$

We define the *domain* and the *image* of a graph as the respective images of π_0 and π_1 over it:

$$\begin{aligned} \text{Dom } R &= \{x \mid (x, y) \in R\} \\ \text{Im } R &= \{y \mid (x, y) \in R\} = \text{Dom } {}^tR \end{aligned}$$

Currying notation

Graphs R can be expressed in curried forms as functors \vec{R} and \overleftarrow{R} :

$$\begin{aligned} \forall x, \quad \vec{R}(x) &= \{y \mid (z, y) \in R \wedge z = x\}. \\ \forall y, \quad \overleftarrow{R}(y) &= \{x \mid (x, z) \in R \wedge z = y\} = {}^t\vec{R}(y). \\ \forall x, y, \quad y \in \vec{R}(x) &\Leftrightarrow (x, y) \in R \Leftrightarrow x \in \overleftarrow{R}(y) \\ \forall x, \quad x \in \text{Dom } R &\Leftrightarrow \vec{R}(x) \neq \emptyset \\ \text{Dom } R \subset E \Rightarrow \text{Im } R &= \bigcup_{x \in E} \vec{R}(x) \wedge \forall y, \overleftarrow{R}(y) = \{x \in E \mid (x, y) \in R\} \end{aligned}$$

They can also appear as functions

$$\begin{aligned} \vec{R}_E &= (E \ni x \mapsto \vec{R}(x)) \\ \overleftarrow{R}_F &= (F \ni y \mapsto \overleftarrow{R}(y)) \\ (\vec{R}) &= \vec{R}_{\text{Dom } R} \\ (\overleftarrow{R}) &= \overleftarrow{R}_{\text{Im } R}. \end{aligned}$$

Functional graphs

The *graph of a function* f is defined by

$$\begin{aligned} \text{Gr } f &= \{(x, f(x)) \mid x \in \text{Dom } f\} \\ \forall x, y, (x, y) \in \text{Gr } f &\Leftrightarrow (x \in \text{Dom } f \wedge y = f(x)) \\ \text{Dom } f &= \text{Dom } \text{Gr } f \\ \text{Im } f &= \text{Im } \text{Gr } f \end{aligned}$$

For any function f and any graph R ,

$$\begin{aligned} \text{Gr } f \subset R &\Leftrightarrow \forall x \in \text{Dom } f, f(x) \in \vec{R}(x) \\ R \subset \text{Gr } f &\Leftrightarrow (\text{Dom } R \subset \text{Dom } f \wedge \forall (x, y) \in R, y = f(x)) \\ R = \text{Gr } f &\Leftrightarrow (\text{Dom } R \subset \text{Dom } f \wedge \forall x \in \text{Dom } f, \vec{R}(x) = \{f(x)\}) \end{aligned}$$

A graph R is *functional* if it is the graph of a function. This condition is equivalently written in either way

$$\begin{aligned} \forall x \in \text{Dom } R, !: \vec{R}(x) \\ \forall x, y \in R, x_0 = y_0 \Rightarrow x_1 = y_1. \end{aligned}$$

It is then the graph of the unique function $\vec{R} = ((\text{Dom } R) \ni x \mapsto \vec{R}(x))$.

For any set E we shall denote $\delta_E = \text{Gr Id}_E$.

Indexed partitions

Two sets E and F are called *disjoint* when $E \cap F = \emptyset$, or equivalently $\forall x \in E, x \notin F$.

A family of sets $(A_i)_{i \in I}$ is called *pairwise disjoint* when $\forall i \neq j \in I, A_i \cap A_j = \emptyset$

For any graph R with $\text{Im } R \subset F$, the family $(\vec{R}(y))_{y \in F}$ is pairwise disjoint if and only if R is functional:

$$(\forall y \neq z \in F, \neg \exists x \in \vec{R}(y), x \in \vec{R}(z)) \Leftrightarrow (\forall (x, y) \in R, \forall z \in F, xRz \Rightarrow y = z)$$

For any function f and any y we define the *fiber* of y under f as

$$f.(y) = \{x \in \text{Dom } f \mid f(x) = y\} = (\overleftarrow{\text{Gr}} f)(y)$$

When $f: E \rightarrow F$ this defines a family $f._F = (f.(y))_{y \in F}$ of pairwise disjoint subsets of E :

$$\begin{aligned} \forall y, z \in F, f.(y) \cap f.(z) \neq \emptyset \Rightarrow \exists x \in f.(y) \cap f.(z), y = f(x) = z \\ \bigcup_{y \in F} f.(y) = E \\ \text{Im } f = \{y \in F \mid f.(y) \neq \emptyset\}. \end{aligned}$$

An *indexed partition* of a set E is a family of nonempty, pairwise disjoint subsets of E , whose union is E .

In other words it is any family of the form $f._ = f._{\text{Im } f}$ for any function f with domain E .

Sum or disjoint union

The binder \coprod of *sum* of any family of sets $(E_i)_{i \in I}$, gives a graph $\coprod_{i \in I} E_i$ defined as

$$\begin{aligned} \coprod_{i \in I} E_i &= \bigcup_{i \in I} \{(i, x) \mid x \in E_i\} \\ \forall i, x, (i, x) \in \coprod_{i \in I} E_i &\Leftrightarrow (i \in I \wedge x \in E_i) \\ (\forall (i, x) \in \coprod_{i \in I} E_i, A(i, x)) &\Leftrightarrow (\forall i \in I, \forall x \in E_i, A(i, x)) \\ (\forall i \in I, E_i \subset E_j) &\Leftrightarrow \coprod_{i \in I} E_i \subset \coprod_{i \in I} E_j \\ E_0 \sqcup \dots \sqcup E_{n-1} &= \coprod_{i \in V_n} E_i \end{aligned}$$

This binder serves as inverse of currying : $R = \coprod_{i \in I} E_i \Leftrightarrow (\text{Dom } R \subset I \wedge \forall i \in I, \vec{R}(i) = E_i)$.

It is also called *disjoint union* as the family of copies $\{(i, x) \mid x \in E_i\}$ of each E_i is pairwise disjoint, with the function from R to I given by π_0 .

Direct and inverse images by a graph

The restriction of a graph R to a set A is defined as

$$R|_A = \{(x, y) \in R \mid x \in A\} = \coprod_{x \in A} \vec{R}(x)$$

The *direct image* of a set A by a graph R is

$$\begin{aligned} R^*(A) &= \text{Im } R|_A = \bigcup_{x \in A} \vec{R}(x) = \{y \mid (x, y) \in R \wedge x \in A\} \subset \text{Im } R \\ \text{Dom } R \subset A &\Leftrightarrow R|_A = R \Rightarrow R^*(A) = \text{Im } R \end{aligned}$$

$$R^*(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} R^*(A_i)$$

$$R^*(\bigcap_{i \in I} A_i) \subset \bigcap_{i \in I} R^*(A_i)$$

$$A \subset B \Rightarrow R^*(A) \subset R^*(B)$$

Similarly, the *inverse image* or *preimage* of a set B by a graph R is

$$R_*(B) = {}^tR^*(B) = \bigcup_{y \in B} \overleftarrow{R}(y) = \{x \mid (x,y) \in R \wedge y \in B\} \subset \text{Dom } R.$$

Direct image and preimage by a function

The direct image of a subset $A \subset \text{Dom } f$ by a function f , denoted $f[A]$ or $f^*(A)$, is

$$f[A] = (\text{Gr } f)^*(A) = \{f(x) \mid x \in A\} \subset \text{Im } f$$

For any $f: E \rightarrow F$ and $B \subset F$, the preimage of B by f , written $f_*(B)$, is defined by

$$f_*(B) = (\text{Gr } f)_*(B) = \{x \in E \mid f(x) \in B\} = \bigcup_{y \in B} f.(y)$$

$$f.(y) = f_*(\{y\})$$

$$f_*(C_F B) = C_E f_*(B)$$

For any family $(B_i)_{i \in I}$ of subsets of F , $f_*(\bigcap_{i \in I} B_i) = \bigcap_{i \in I} f_*(B_i)$ where intersections are respectively interpreted as subsets of F and E .

2.7. Products and powerset

Cartesian product of two sets

For any two sets E and F , the (cartesian) *product* $E \times F$ is the set of (x,y) where $x \in E$ and $y \in F$.

$$\begin{aligned} E \times F &= \prod_{x \in E} F \\ \prod_{i \in I} E_i &= \bigcup_{i \in I} \{i\} \times E_i \subset I \times \bigcup_{i \in I} E_i \\ (E \subset E \wedge F \subset F) &\Rightarrow E \times F \subset E \times F \\ E \times \emptyset &= \emptyset = \emptyset \times E \end{aligned}$$

For any graph R ,

$$\begin{aligned} R \subset E \times F &\Leftrightarrow (\text{Dom } R \subset E \wedge \text{Im } R \subset F) \\ R \downarrow_E &= R \cap (E \times \text{Im } R) \end{aligned}$$

Quantifiers over a product have equivalent curried views:

$$\begin{aligned} (\exists (x,y) \in E \times F, A(x,y)) &\Leftrightarrow (\exists x \in E, \exists y \in F, A(x,y)) \Leftrightarrow (\exists y \in F, \exists x \in E, A(x,y)) \\ (\forall (x,y) \in E \times F, A(x,y)) &\Leftrightarrow (\forall x \in E, \forall y \in F, A(x,y)) \Leftrightarrow (\forall y \in F, \forall x \in E, A(x,y)) \\ (\exists x \in E, A(x) \vee B(x)) &\Leftrightarrow ((\exists x \in E, A(x)) \vee (\exists x \in E, B(x))) \\ (\forall x \in E, A(x) \wedge B(x)) &\Leftrightarrow ((\forall x \in E, A(x)) \wedge (\forall x \in E, B(x))) \\ (\exists x \in E, C \vee A(x)) &\Leftrightarrow ((C \wedge E \neq \emptyset) \vee \exists x \in E, A(x)) \\ (\forall x \in E, C \wedge A(x)) &\Leftrightarrow ((C \vee E = \emptyset) \wedge \forall x \in E, A(x)) \end{aligned}$$

Quantifiers $(\forall x,y \in E,)$ can be equivalently written $(\forall (x,y) \in E \times E,)$ while the notation $\forall x \neq y \in E$ can be read as one quantifier with domain $\{(x,y) \in E \times E \mid x \neq y\} = (E \times E) \setminus \delta_E$.

Finite products, operations and relations

Beyond the binary operator of cartesian product, there is an n -ary product operator between any number $n > 1$ of sets : $E_0 \times \dots \times E_{n-1}$ is the set of all n -tuples (x_0, \dots, x_{n-1}) such

that $\forall i \in V_n, x_i \in E_i$.

For example, the ternary product is

$$E \times F \times G = \bigcup_{x \in E} \bigcup_{y \in F} \{(x, y, z) \mid z \in G\} = \{(x, y, z) \mid ((x, y), z) \in (E \times F) \times G\}$$

where the latter notation abbreviates the use of the [evaluators for this representation of triples](#). It can also be seen as directly justified by the set generation principle:

$$(\forall (x, y, z) \in E \times F \times G, A(x, y, z)) \Leftrightarrow (\forall x \in E, \forall y \in F, \forall z \in G, A(x, y, z))$$

This finally justifies the most natural formalization of n -ary operations, as functions with domain a product of n sets, with [evaluator given in 2.3](#). It justifies our first [notation of 2.3 for operation definers](#), as a use of the abbreviated notation of 2.6 for binders on graphs, extensible to any arity.

The notion of n -ary relation R will usually be formalized as that of a set of n -tuples, subset of the product of its domains. This way, a binary relation between sets E and F will be a graph $R \subset E \times F$. The equality relation on a set E is so formalized as δ_E , called the *diagonal* of $E \times E$.

For a formalization which specifies the domains E and F of arguments, we can take the triple (E, F, R) .

Translating operators into predicates

The use of any functor symbol K in a generic theory is equivalent to that of a binary predicate symbol R with the functionality axiom

$$\forall x, \exists! y, xRy$$

It is meant as the "graph of K ", i.e. defined from K by

$$xRy \Leftrightarrow y = K(x)$$

Terms of first-order logic using K cannot be re-expressed using R , but formulas can: each formula with each occurrence of K in it, can be written $A(K(x))$ where x abbreviates a term and A abbreviates a formula as a unary predicate; then equivalently replaced by either of

$$\begin{aligned} \exists y, xRy \wedge A(y) \\ \forall y, xRy = A(y) \end{aligned}$$

The same goes for other arities : the role of any n -ary operator can be played by an $(n+1)$ -ary predicate with the functionality axiom. This generalization can be explained as replacing the argument by a tuple.

This construction does not fit for the operators of our set theory, due to its use of open quantifiers. Yet it works when domains are sets: n -ary operations can be represented as functional $(n+1)$ -ary graphs.

More primitive symbols

Let us [strengthen set theory](#) by 3 extensions that declare more kinds of classes [to be sets](#) : powersets, exponentiations and products. These extensions are equivalent, in the sense that accepting anyone of them as primitive suffices to get both others as developments from it. They are similar to those provided by the [set generation principle](#), but no more fit its condition ; this lack of justification has deep connections with [incompleteness aspects of mathematics](#).

To fit our way of formalizing set theory (unlike the ZF way introduced below), each such extension is made of

- A symbol K (namely, two operators and one binder) naming as a set the given class C ,

with arguments the parameters of C here abbreviated as a family y ranging over the intended [definiteness class](#) of K ;

- The statement $\forall_{dK y, \text{Set } K(y)} \wedge \forall x, x \in K(y) \Rightarrow C_y(x)$.

Powerset. $\forall_{\text{Set } E, \text{Set } \wp(E)} \wedge \forall F, F \in \wp(E) \Rightarrow (\text{Set } F \wedge F \subset E)$

We shall abbreviate $\in \wp$ as \subset also in binders: for example $(\forall F \subset E, \dots) \Leftrightarrow (\forall F \in \wp(E), \dots)$.

Exponentiation. $\forall_{\text{Set } E, F, \text{Set } F^E} \wedge \forall f, f \in F^E \Rightarrow f: E \rightarrow F$

Product of a family of sets. The binder \prod generalizes the finite product operators to any family of sets $(\forall i \in I, \text{Set } E_i)$:

$$\forall x, x \in \prod_{i \in I} E_i \Leftrightarrow (\text{Fnc}(x) \wedge \text{Dom } x = I \wedge \forall i \in I, x_i \in E_i).$$

Like with [tuples](#), the function evaluator carried in the reverse way defines a family of *projections* π_i (keeping implicit their dependence on the given product which is usually fixed in context):

$$\forall i \in I, \pi_i: \prod_{j \in I} E_j \rightarrow E_i.$$

These symbols preserve inclusion as follows

$$\begin{aligned} F \subset E &\Rightarrow \wp(F) \subset \wp(E) \\ F \subset E &\Rightarrow F^G \subset E^G \\ (\forall i \in I, F_i \subset E_i) &\Rightarrow \prod_{i \in I} F_i \subset \prod_{i \in I} E_i \end{aligned}$$

Their equivalence

From any of these 3 symbols, both others can be defined as follows.

$$\begin{aligned} \prod_{i \in I} E_i &= \{x \in (\bigcup_{i \in I} E_i)^I \mid \forall i \in I, x_i \in E_i\} = \{\vec{R} \mid R \subset \prod_{i \in I} E_i \wedge \forall i \in I, \exists!: \vec{R}(i)\} \\ F^E &= \prod_{x \in E} F = \{\vec{R} \mid R \subset E \times F \wedge \forall x \in E, \exists!: \vec{R}(x)\} \\ \wp(E) &= \{f \cdot (1) \mid f \in V_2^E\} \end{aligned}$$

Even some cases are expressible from previous tools:

$$\begin{aligned} F^{\{a\}} &= \{\{a\} \ni x \mapsto y \mid y \in F\} \\ F^\emptyset &= \prod \emptyset \mapsto = \{\emptyset \mapsto\} \\ \wp(\{a\}) &= \{\emptyset, \{a\}\} \\ \prod_{i \in I \cup J} E_i &= \{(i \in I? x_i : y_i)_{i \in I \cup J} \mid (x, y) \in \prod_{i \in I} E_i \times \prod_{i \in J} E_i\} \\ (\exists i \in I, E_i = \emptyset) &\Rightarrow \prod_{i \in I} E_i = \emptyset \\ (\forall i \in I, \exists!: E_i) &\Rightarrow \prod_{i \in I} E_i = \{(1 E_i)_{i \in I}\} \end{aligned}$$

Cantor's Theorem

This famous theorem is simply expressible as

$$\forall_{\text{Fnc } f, \wp(\text{Dom } f) \not\subset \text{Im } f.$$

Proof for the case when all $f(x)$ are sets (to which the general case is reducible):

$$(E = \text{Dom } f \wedge F = \{x \in E \mid x \notin f(x)\}) \Rightarrow (\forall x \in E, x \in F \neq x \in f(x)) \Rightarrow (\forall x \in E, F \neq f(x)) \Rightarrow F \notin \text{Im } f. \blacksquare$$

[Russell's paradox](#) may be seen as the particular case of Cantor's theorem for $f = \text{Id}_E$.

The ZF approach

In the traditional ZF formalism for set theory, each such extension is formalized as first made of a mere statement without any special symbol, keeping \in as the only primitive structure:

$$\forall \dots y, \exists X, \forall x, x \in X \Leftrightarrow C_y(x).$$

Namely, the powerset axiom (existence of the powerset) is written $\forall E, \exists X, \forall F, F \in X \Leftrightarrow F \subset E$, from which the statements of existence of exponentiations and products come as theorems.

Then, the uses of each symbol K with argument y as above, naming these sets, can be justified as developed from there as abbreviations by the following rules.

- $\forall x \in K$ means $\forall x, C(x) \Rightarrow$;
- The equality $(X=K)$ means $(\forall x, x \in X \Leftrightarrow C(x))$ or equivalently $(\forall x \in X, C(x)) \wedge (\forall_C x, x \in X)$;
- Any other formulas using K with y free, written as $A(K)$, mean the formula so abbreviated as $\exists X, (X=K) \wedge A(X)$.
- To justify binding y on terms using K , involves the axiom schema of replacement.

But our framework does not even allow the first of these cases, because it involves an open quantifier, which cannot be replaced by a bounded formula because these extensions do not fit the condition of the set generation principle.

This impossibility to interpret the equality formula $(X=K)$ means that even a given set identical to a class C , would not be recognizable as such. Hence the need to accept the symbols K as primitive, for giving effective sense to the statement that C coincides with a set.

2.8. Injections, bijections

Composition, restriction

The *composite* of any 2 functions f, g such that $\text{Im } f \subset \text{Dom } g$ is defined by

$$\begin{aligned} g \circ f &= ((\text{Dom } f) \ni x \mapsto g(f(x))) \\ (f: E \rightarrow F \wedge g: F \rightarrow G) &\Rightarrow g \circ f: E \rightarrow G \end{aligned}$$

(This notation will be abusively kept if g is a functor, in which case this is a schema of definitions).

Similarly with more than 2 functions: adding a third one $h: G \rightarrow H$ to the above conditions, we define

$$h \circ g \circ f = (h \circ g) \circ f = h \circ (g \circ f) = (E \ni x \mapsto h(g(f(x))))$$

The *restriction* of a function f to a set $A \subset \text{Dom } f$ is

$$\begin{aligned} f|_A &= (A \ni x \mapsto f(x)) = f \circ \text{Id}_A. \\ \text{Gr}(f|_A) &= (\text{Gr } f)|_A \\ f[A] &= \text{Im}(f|_A) \\ \forall_{\text{Fnc}} f, g, \text{Gr } g \subset \text{Gr } f &\Leftrightarrow (\text{Dom } g \subset \text{Dom } f \wedge g = f|_{\text{Dom } g}) \end{aligned}$$

When g is a restriction of f , we also say that f is an *extension* of g .

Injections, bijections, inverse

A function f is called *injective* (or : an *injection*), if ${}^t\text{Gr } f$ is a functional graph:

$$\begin{aligned} \text{Inj } f &\Leftrightarrow (\forall x \neq x' \in \text{Dom } f, f(x) \neq f(x')) \\ &\Leftrightarrow (\forall x, x' \in \text{Dom } f, f(x) = f(x') \Rightarrow x = x') \\ (f: E \hookrightarrow F) &\Leftrightarrow (f: E \rightarrow F \wedge \text{Inj } f) \\ \text{Im } f \subset F &\Leftrightarrow (\text{Inj } f \Leftrightarrow \forall y \in F, \exists! f.(y)) \end{aligned}$$

The *inverse* of any injection f , is the function f^{-1} with graph ${}^t\text{Gr } f$:

$$\begin{aligned} f^{-1} &= {}_1f. = (\text{Im } f \ni y \mapsto {}_1f.(y)) \\ \text{Gr}(f^{-1}) &= {}^t\text{Gr } f \\ (f^{-1})^{-1} &= f. \end{aligned}$$

A function $f: E \rightarrow F$ is said *bijective* (or a *bijection*) from E onto F , if it is both injective and surjective :

$$\begin{aligned} f: E \leftrightarrow F &\Leftrightarrow (f: E \rightarrow F \wedge \forall y \in F, \exists! f.(y)) \\ &\Leftrightarrow (f: E \hookrightarrow F \wedge \text{Im } f = F) \\ &\Leftrightarrow (f: E \rightarrow F \wedge \text{Inj } f) \\ &\Rightarrow f^{-1}: F \leftrightarrow E. \end{aligned}$$

In particular $\text{Id}_E: E \leftrightarrow E$.

Diverse properties

Proposition. Let $f: E \rightarrow F$ and $g: F \rightarrow G$. Then

1. $(\text{Inj } f \wedge \text{Inj } g) \Rightarrow \text{Inj}(g \circ f)$
2. $\text{Inj}(g \circ f) \Rightarrow \text{Inj } f$
3. $\text{Im}(g \circ f) = g[\text{Im } f] \subset \text{Im } g$.

Proofs:

1. $\forall x, y \in E, g(f(x)) = g(f(y)) \Rightarrow f(x) = f(y) \Rightarrow x = y$.
2. $\forall x, y \in E, f(x) = f(y) \Rightarrow g(f(x)) = (g \circ f)(x) = (g \circ f)(y) \Rightarrow x = y$.
3. $\forall z \in G, z \in \text{Im}(g \circ f) \Leftrightarrow (\exists x \in E, g(f(x)) = z) \Leftrightarrow (\exists y \in \text{Im } f, g(y) = z) \Leftrightarrow z \in g[\text{Im } f]$. ■

From 3. we can obviously deduce

$$\begin{aligned} \text{Im}(g \circ f) = G &\Rightarrow \text{Im } g = G \\ \text{Im } f = F &\Rightarrow \text{Im}(g \circ f) = \text{Im } g \\ (f: E \rightarrow F) &\Rightarrow (g: F \rightarrow G \Leftrightarrow g \circ f: E \rightarrow G) \end{aligned}$$

Let $f: E \rightarrow F$, and a function g with $\text{Dom } g = F$. Then, easily,

$$\begin{aligned} g \circ f = \text{Id}_E &\Leftrightarrow (\forall x \in E, \forall y \in F, f(x) = y \Rightarrow g(y) = x) \\ &\Leftrightarrow \text{Gr } f \subset {}^t\text{Gr } g \Leftrightarrow (\forall y \in F, f.(y) \subset \{g(y)\}) \\ &\Leftrightarrow (\forall x \in E, f(x) \in g.(x)) \Leftrightarrow (\text{Inj } f \wedge g|_{\text{Im } f} = f^{-1}) \\ &\Rightarrow (\text{Im } f = F \Leftrightarrow g = f^{-1} \Leftrightarrow (\text{Inj } g \wedge \text{Im } g \subset E)) \end{aligned}$$

Proof of the non-obvious step :

$$(g \circ f = \text{Id}_E \wedge \text{Inj } g \wedge \text{Im } g \subset E) \Rightarrow (\forall y \in F, g(f(g(y))) = g(y) \therefore y = f(g(y)) \in \text{Im } f)$$

Proposition. $(f: E \rightarrow F \wedge g: F \rightarrow E \wedge g \circ f = \text{Id}_E) \Rightarrow (\text{Im } f = F \Leftrightarrow \text{Inj } g \Leftrightarrow f \circ g = \text{Id}_F \Leftrightarrow g = f^{-1})$.

Proof : $(\text{Inj } g \wedge g \circ f \circ g = g \circ \text{Id}_F) \Rightarrow f \circ g = \text{Id}_F$. ■

Proposition. 1) If f, g are injective and $\text{Im } f = \text{Dom } g$, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
 2) If $f, h : E \rightarrow F$ and $g : F \rightarrow E$ then $(g \circ f = \text{Id}_E \wedge h \circ g = \text{Id}_F) \Leftrightarrow ((g : F \hookrightarrow E) \wedge f = h = g^{-1})$.

Proofs:

1) $\forall x \in \text{Dom } f, \forall y \in \text{Im } g, (g \circ f)(x) = y \Leftrightarrow f(x) = g^{-1}(y) \Leftrightarrow x = (f^{-1} \circ g^{-1})(y)$.

Other method: $(g \circ f)^{-1} = (g \circ f)^{-1} \circ g \circ g^{-1} = (g \circ f)^{-1} \circ g \circ f^{-1} \circ g^{-1} = f^{-1} \circ g^{-1}$.

2) We deduce $f = h$ from $f = h \circ g \circ f = h$, or from $\text{Gr } f \subset {}^t\text{Gr } g \subset \text{Gr } h$. The rest is obvious. ■

Canonical functions

For any objects x, y we shall say that x determines y if there is an invariant functor T (known but kept implicit) such that $T(x) = y$. Then the role of y as free variable can be played by the term $T(x)$, thus by the use of x .

Similarly, a function $f : E \rightarrow F$ will be said *canonical* if it is defined as $E \ni x \mapsto T(x)$ where T is some invariant functor.

Examples of important canonical injections :

- $E \subset F \Rightarrow \text{Id}_E : E \hookrightarrow F$
- $(x \mapsto \{x\}) : E \hookrightarrow \wp(E)$
- $\text{Gr} : F^E \hookrightarrow \wp(E \times F)$.

A bijection f will be said *bicanonical* if both f and f^{-1} are canonical.

When a bijection $f : E \leftrightarrow F$ is canonical (resp. bicanonical), we shall write $f : E \rightleftharpoons F$ (resp. $f : E \rightleftharpoons F$); or, with its defining functor φ as $\varphi : E \rightleftharpoons F$ which means that $(E \ni x \mapsto \varphi(x))$ is injective with image F . We shall write $E \rightleftharpoons F$ (resp. $E \rightleftharpoons F$) to mean the existence of a canonical (resp. bicanonical) bijection, that is kept implicit. These will often look like equality properties on numbers, as the existence of a bijection between finite sets implies the equality of their numbers of elements.

Canonical bijections can fail to be bicanonical especially when their defining functor is not injective:

$$\begin{aligned} V_2^E &= \wp(E) \\ E \neq \emptyset &= \{x\}^E = \{x\} \\ E \times \{x\} &= E \\ E \times F &= F \times E \\ E \times \{x\} &= \{x\} \times E = E^{\{x\}} \\ E \times \{0\} &= E. \end{aligned}$$

This meta-relation between sets is preserved by constructions: if $E \rightleftharpoons E'$ and $F \rightleftharpoons F'$ then $\wp(E) \rightleftharpoons \wp(E')$, $E \times F \rightleftharpoons E' \times F'$ and $F^E \rightleftharpoons F'^{E'}$ (using the direct image of the graph), and the same for \rightleftharpoons .

This is how transposition ($E \times F \rightleftharpoons F \times E$) extends to both graphs and operations:

$$\begin{aligned} \wp(E \times F) &= \wp(F \times E) \\ G^{E \times F} &= G^{F \times E} \\ \forall f \in G^{E \times F}, {}^t f &= (F \times E \ni (x, y) \mapsto f(y, x)) \in G^{F \times E}. \end{aligned}$$

Sums of functions

The binder \coprod of [sum of sets](#) defines canonical bijections (whose inverses, defined by $R \mapsto \vec{R}_I$, depend on I):

$$\begin{aligned} (\wp(E))^I &= \wp(I \times E) \\ \prod_{i \in I} \wp(A_i) &= \wp(\prod_{i \in I} A_i) \end{aligned}$$

Any sum of sets $S = \coprod_{i \in I} A_i$ has its family of *natural injections* j_i given by the curried ordered pair definer :

$$\forall i \in I, j_i : A_i \hookrightarrow S.$$

The sum over I of functions f_i where $\forall i \in I, A_i = \text{Dom } f_i$ is defined by

$$\begin{aligned} \coprod_{i \in I} f_i &= (S \ni (i, x) \mapsto f_i(x)) \\ g &= \coprod_{i \in I} f_i \Leftrightarrow (\text{Dom } g = S \wedge \forall i \in I, f_i = g \circ j_i = \vec{g}(i)) \end{aligned}$$

It is the inverse of the currying of binary operations, which we shall denote like for graphs: for any sets E, F and any operation (function) B with domain $E \times F \neq \emptyset$ (from which E and F can be restored as $E = \text{Dom } \text{Dom } B$ and $F = \text{Im } \text{Dom } B$, while if $E \times F = \emptyset$ the notation can be abusively kept while taking E and F from context),

$$\begin{aligned} \vec{B} &= (E \ni x \mapsto (F \ni y \mapsto B(x, y))) \\ \overleftarrow{B} &= {}^t\vec{B} = (F \ni y \mapsto (E \ni x \mapsto B(x, y))) \end{aligned}$$

Hence the canonical bijections (bicanonical if $I = \text{Dom } S$, thus if $E \neq \emptyset$ in the case $I \times E$)

$$\begin{aligned} \forall_{\text{Set}} F, \prod_{i \in I} F^{A_i} &= F^S \\ \forall_{\text{Set}} E, F, (F^E)^I &= F^{I \times E} \\ (\forall F : S \rightarrow \text{Set}) \prod_{i \in I} \prod_{x \in A_i} F_{(i, x)} &= \prod_{c \in S} F_c \\ (E \times F) \times G &= E \times F \times G \\ f \in (F^E)^I &= \prod_{i \in I} \text{Gr } f_i \subset \wp(I \times (E \times F)) = \wp((I \times E) \times F) \supset \text{Gr } \prod_{i \in I} f_i \end{aligned}$$

Product of functions or recurring

Both curried forms \vec{R} and \overleftarrow{R} of graphs $R \subset E \times F$ are exchanged by a bijection $\wp(F)^E \leftrightarrow \wp(E)^F$ defined with parameter F .

The similar bijection between both curried forms of operations with a given domain $I \times E$, is defined by the binder \sqcap of *product of functions* all with domain E , and indexed by I :

$$\begin{aligned} (\forall i \in I, \text{Dom } f_i = E) &\Rightarrow \sqcap_{i \in I} f_i = (E \ni x \mapsto (f_i(x))_{i \in I}) \\ \forall g, g = \sqcap_{i \in I} f_i &\Leftrightarrow (g : E \rightarrow \prod_{i \in I} \text{Im } f_i \wedge \forall i \in I, f_i = \pi_i \circ g) \\ (F : I \rightarrow \text{Set}) &\Rightarrow \sqcap : \prod_{i \in I} F_i^E = (\prod_{i \in I} F_i)^E \\ \text{Set } F &\Rightarrow \sqcap : (F^E)^I = (F^I)^E \end{aligned}$$

These bijections are canonical for nonempty families ($I \neq \emptyset$) as only these determine E , and are usually bicanonical as their inverse uses I which is definable when $E \neq \emptyset$; anyway we shall abusively keep the notation \Rightarrow for the general case while E remains fixed by the context.

A particular case is the binary product ($I = V_2$):

$$\begin{aligned} \forall_{\text{Fnc}} f, g, \text{Dom } f = \text{Dom } g = E &\Rightarrow f \sqcap g = (E \ni x \mapsto (f(x), g(x))) \\ I^{E \times F^E} &= (I \times F)^E \\ (\prod_{i \in I} F_i)^E &= \prod_{\phi \in I^E} \prod_{x \in E} F_{\phi(x)} \\ \forall_{\text{Fnc}} f, \text{Gr } f &= \text{Im}(\text{Id}_{\text{Dom } f} \sqcap f) \end{aligned}$$

2.9. Binary relations on a set

A *binary relation on a set* E is a relation with both domains equal to E , thus formalized by a graph $R \subset E \times E$. Let us abbreviate $E \times E$ as E^2 and $(x, y) \in R$ as $x R y$

Preimages and products

For any function $f: E \rightarrow F$ and any binary relation R on F , the *preimage* of R by f is the binary relation $f_*(R)$ on E defined as

$$\{(x, y) \in E^2 \mid f(x) R f(y)\}$$

For any family of a binary relations $(R_i)_{i \in I}$ on respective sets F_i , their *product* is the binary relation on $P = \prod_{i \in I} F_i$ defined as

$$\{(x, y) \in P^2 \mid \forall i \in I, x_i R_i y_i\} = \bigcap_{i \in I} \Pi_{i*}(R_i).$$

(It is actually in canonical bijection with the product of sets $\prod_{i \in I} R_i$)

In particular with a constant family : any binary relation on a set F defines one on any F^I .

Both operations can be combined in one step: for any family of functions $f_i: E \rightarrow F_i$, the relation

$$\{(x, y) \in E^2 \mid \forall i \in I, f_i(x) R_i f_i(y)\} = \bigcap_{i \in I} f_{i*}(R_i)$$

is the preimage of the product relation of the R_i on P , by the function $\prod_{i \in I} f_i: E \rightarrow P$.

These concepts will be later generalized to other first-order structures.

Basic properties

A binary relation R on a set E will be said

$$\text{reflexive} \Leftrightarrow \forall x \in E, x R x \Leftrightarrow \delta_E \subset R \Rightarrow \text{Dom } R = \text{Im } R = E$$

$$\text{irreflexive} \Leftrightarrow \forall x \in E, \neg(x R x)$$

$$\text{symmetric} \Leftrightarrow (\forall x, y \in E, x R y \Rightarrow y R x) \Leftrightarrow R \subset {}^t R \Leftrightarrow R = {}^t R \Leftrightarrow (\vec{R}) = (\overleftarrow{R})$$

$$\text{antisymmetric} \Leftrightarrow \forall x, y \in E, (x R y \wedge y R x) \Rightarrow x = y$$

$$\text{transitive} \Leftrightarrow \forall x, y, z \in E, (x R y \wedge y R z) \Rightarrow x R z$$

These properties are preserved by transposition, preimages and products, except that the preimage of an antisymmetric relation by a non-injective function may not be antisymmetric, and the empty product ($I = \emptyset$) is not irreflexive.

In particular for $F \subset E$, the restriction of a relation $R \subset E \times E$ to F , namely $R \cap (F \times F)$, is its preimage by $\text{Id}_F: F \hookrightarrow E$, thus preserves all properties.

Example. Let $A \subset E^E$ and $R = \bigcup_{f \in A} \text{Gr } f$. Then

$$(\text{Id}_E \in A) \Rightarrow R \text{ is reflexive}$$

$$(\forall f, g \in A, g \circ f \in A) \Rightarrow R \text{ is transitive}$$

$$(\forall f \in A, f: E \rightarrow E \wedge f^{-1} \in A) \Rightarrow R \text{ is symmetric}$$

For any transitive binary relation R we denote $x R y R z \Leftrightarrow (x R y \wedge y R z) \Rightarrow x R z$.

Preorders and orders

A *preorder* is a reflexive and transitive binary relation. An *order* is an antisymmetric preorder.

A *preordered set* is (an ordered pair of) a set with a chosen preorder on it.

An *ordered set* is a set with a chosen order, usually written as \leq or \leq_E . The formula $x \leq y$ can be read « x is less than y », or « y is greater than x ».

Two elements x, y of an ordered set are said *comparable* when $(x \leq y \vee y \leq x)$.

On the class of sets, the meta-relation \Rightarrow (existence of a canonical bijection) is a meta-preorder.

The Boolean pair V_2 is naturally ordered by \Rightarrow which coincides there with the usual order between numbers.

The resulting product order on $V_2^E = \wp(E)$ (and thus any set of sets) is that of inclusion (\subset).

Any relation $R \subset E \times F$ defines a preorder on E as the preimage by \vec{R} of the inclusion order:

$$\{(x, y) \in E^2 \mid \vec{R}(x) \subset \vec{R}(y)\}.$$

Conversely, any preorder R on a set E is so definable from itself in both ways (exchanged by transposition):

$$\forall x, y \in E, \vec{R}(y) \subset \vec{R}(x) \Leftrightarrow x R y \Leftrightarrow \overleftarrow{R}(x) \subset \overleftarrow{R}(y)$$

Proof. Transitivity is equivalent to $(\forall x, y \in E, x R y \Rightarrow \vec{R}(y) \subset \vec{R}(x))$ and to $(\forall x, y \in E, x R y \Rightarrow \overleftarrow{R}(x) \subset \overleftarrow{R}(y))$.

Reflexivity is equivalent to the converses:

$$\forall x \in E, (\forall y \in E, \vec{R}(x) \subset \vec{R}(y) \Rightarrow y R x) \Leftrightarrow x R x \Leftrightarrow (\forall y \in E, \overleftarrow{R}(x) \subset \overleftarrow{R}(y) \Rightarrow x R y). \blacksquare$$

Thus R is the preimage of \subset by (\vec{R}) . Moreover,

$$\forall x, y \in E, \overleftarrow{R}(x) = \overleftarrow{R}(y) \Leftrightarrow (\overleftarrow{R}(x) \subset \overleftarrow{R}(y) \wedge \overleftarrow{R}(y) \subset \overleftarrow{R}(x)) \Leftrightarrow (x R y \wedge y R x)$$

so that (\vec{R}) is injective if and only if R is an order.

The idea of considering only one order written \leq on each set E , may be formalized by replacing the ordered set (E, R) by $(\text{Im}(\vec{R}), \subset)$, though this does not exactly fit the following convention : any subset of an ordered set will be implicitly also ordered by restriction.

Strict and total orders

A *strict order* is a binary relation both transitive and irreflexive; and thus also antisymmetric.

Strict orders $<$ bijectively correspond to orders \leq by

$$\begin{aligned} x < y &\Leftrightarrow (x \leq y \wedge x \neq y). \\ x \leq y &\Leftrightarrow (x < y \vee x = y). \end{aligned}$$

A *total order* on a set E is an order where all elements are comparable $(\forall x, y \in E, x \leq y \vee y \leq x)$, i.e. $(\leq \cup \leq^t) = E \times E$.

Equivalent definitions:

- an order \leq related with its strict order $<$ by $\forall x, y \in E, x < y \Leftrightarrow y \leq x$.
- a transitive relation \leq such that $\forall x, y \in E, x \leq y \Leftrightarrow (y \leq x \Rightarrow x = y)$.

A *strict total order* on a set E is a strict order associated with a total order, i.e. any transitive relation $<$ such that, equivalently

$$\begin{aligned}\forall x, y \in E, x < y &\Leftrightarrow (y < x \vee x = y) \\ \forall x, y \in E, x = y &\Leftrightarrow (y < x \vee x < y)\end{aligned}$$

Equivalence relations

An *equivalence relation* is a symmetric preorder.

On any set E , the equality relation δ_E is an equivalence relation (as already expressed by the [axioms of equality](#)), and the only relation altogether symmetric, antisymmetric and reflexive.

On any product of sets, the product of equality relations is also equality.

The preimage of equality by any $f: E \rightarrow F$ is an equivalence relation on E we shall write \sim_f :

$$\begin{aligned}\sim_f &= \{(x, y) \in E \mid f(x) = f(y)\} = \text{II}(f \circ f) \\ \text{Inj } f &\Leftrightarrow \sim_f = \delta_E\end{aligned}$$

Conversely, any equivalence relation R is so definable from itself : $R = \sim_{\vec{R}}$. Indeed

$$\forall x, y \in E, x R y \Leftrightarrow \vec{R}(y) \subset \vec{R}(x) \Leftrightarrow \vec{R}(x) \subset \vec{R}(y) \Leftrightarrow \vec{R}(x) = \vec{R}(y)$$

Quotient functions

Proposition. $\forall_{\text{Set}} E, F, G, \forall f, (f: E \rightarrow F) \Rightarrow (G^F \ni g \mapsto g \circ f) : G^F \leftrightarrow \{h \in G^E \mid \sim_f \subset \sim_h\}$.

Proof : $\forall \dots \forall f \in F^E, \forall h \in G^E, H = \text{Im}(f \times h) \Rightarrow$

$$\begin{aligned}\sim_f \subset \sim_h &\Leftrightarrow (\forall x, x' \in E, f(x) = f(x') \Rightarrow h(x) = h(x')) \Leftrightarrow (\forall (y, z), (y', z') \in H, y = y' \Rightarrow z = z') \\ \text{Dom } H &= \text{Im } f \wedge \forall g \in G^F, h = g \circ f \Leftrightarrow H \subset \text{Gr } g. \blacksquare\end{aligned}$$

For any functions f, h such that $\text{Dom } f = \text{Dom } h \wedge \sim_f \subset \sim_h$, this unique $g: \text{Im } f \rightarrow \text{Im } h$ such that $h = g \circ f$ will be called the *quotient* of h by f :

$$\begin{aligned}h/f &= (\text{Im } f \ni y \mapsto \imath h[f \cdot (y)]) \\ \text{Gr } (h/f) &= \text{Im } (f \times h) \\ \text{Inj } f &\Leftrightarrow (f^{-1} = \text{Id}_{\text{Dom } f} / f \wedge h/f = h \circ f^{-1}) \\ \sim_f = \sim_h &\Leftrightarrow \text{Inj}(h/f) \Leftrightarrow (h/f)^{-1} = f/h\end{aligned}$$

In the more general case $\text{Im } f \subset \text{Dom } g$ (without surjectivity of f),

$$\begin{aligned}h = g \circ f &\Leftrightarrow (\sim_f \subset \sim_h \wedge h/f = g|_{\text{Im } f}) \\ &\Leftrightarrow (\sim_f = \sim_h \Rightarrow f = (g|_{\text{Im } f})^{-1} \circ h)\end{aligned}$$

The formula that \sim_f is an equivalence relation,

$$\forall x, y \in \text{Dom } f, x \sim_f y \Leftrightarrow \vec{f}(x) = \vec{f}(y)$$

where $(\vec{f}) = f \circ f$, reflects the injectivity of $(f \cdot)$ which can be directly seen as

$$\forall y \in \text{Im } f, \forall z, f \cdot (y) = f \cdot (z) \Rightarrow f \cdot (y) \cap f \cdot (z) = f \cdot (y) \neq \emptyset \Rightarrow y = z.$$

Partitions

A *partition* of a set E is a set of nonempty, pairwise disjoint sets, whose union is E . In other words, it is a set P of subsets of E such that Id_P is an [indexed partition](#) of E . The image of any indexed partition is a partition of the same set: for any function f with image E , $\text{Im}(f \cdot) = f.[E] = \text{Im}(f \circ f)$ is a partition of E .

$\forall R \subset E \times E$, if $P = \vec{R}[E]$ then

$$(\forall x, y \in E, x \in \vec{R}(y) \Leftrightarrow \vec{R}(x) = \vec{R}(y)) \Leftrightarrow (\forall x \in E, \forall A \in P, x \in A \Leftrightarrow \vec{R}(x) = A) \Leftrightarrow \text{Id}_P = (\vec{R}_{E \cdot})$$

Thus if R is an equivalence relation then P is a partition.

Inversely, any partition P of E defines a function g and then a graph R by

$$\begin{aligned} \text{Dom } g &= E \wedge (g \cdot) = \text{Id}_P \therefore P = \text{Dom } (g \cdot) = \text{Im } g \\ R &= \coprod [g \subset E \times E \therefore g = \vec{R}_E \therefore (P = \vec{R}[E] \wedge \text{Id}_P = (\vec{R}_{E \cdot})) \end{aligned}$$

thus R is an equivalence relation on E and

$$\forall x, y \in E, x R y \Leftrightarrow y \in \imath \{A \in P \mid x \in A\} \Leftrightarrow (\exists A \in P, x \in A \wedge y \in A) \Leftrightarrow (\forall A \in P, x \in A \Rightarrow y \in A).$$

For any equivalence relation R on E , the partition $\vec{R}[E] = \text{Im } (\vec{R})$ is called the *quotient of E by R* , and written E/R , so that

$$(\vec{R}) : E \rightarrow E/R.$$

For any $x \in E$, the set $\vec{R}(x) = \imath \{A \in E/R \mid x \in A\}$ is called the *class of x by R* .

For any function f such that $\text{Dom } f = E \wedge R \subset \sim_f$ let us abbreviate $f(\vec{R})$ as f/R :

$$\begin{aligned} f/R &: E/R \rightarrow \text{Im } f \\ \text{Inj } f/R &\Leftrightarrow R = \sim_f \end{aligned}$$

So when $R = \sim_f$ the role of E/R may be played by $\text{Im } f$.

In particular, any preorder R on a set E gives an equivalence relation $R \cap {}^t R$, and the role of $E/(R \cap {}^t R)$ may be played by $\text{Im } (\vec{R})$.

2.10. Axiom of choice

Properties of curried composition

Proposition. $\forall_{\text{Set}} E, F, G, \forall f \in F^E$,

1. $\text{Im } f = F \Rightarrow \text{Inj}(G^F \ni g \mapsto g \circ f)$
2. $(\text{Inj}(G^F \ni g \mapsto g \circ f) \wedge \exists^{\geq 2}: G) \Rightarrow \text{Im } f = F$
3. $(\text{Inj } f \wedge G \neq \emptyset) \Rightarrow \{g \circ f \mid g \in G^F\} = G^E$
4. $(\{g \circ f \mid g \in G^F\} = G^E \wedge \exists^{\geq 2}: G) \Rightarrow \text{Inj } f$
5. $(E \subset F \wedge G \neq \emptyset) \Rightarrow \{g|_E \mid g \in G^F\} = G^E$
6. $(\text{Inj } f \wedge E \neq \emptyset) \Rightarrow \exists g \in E^F, g \circ f = \text{Id}_E$

Proofs :

1. done in 2.9 ; simpler proof: $\forall g, h \in G^F, (\forall x \in E, g(f(x)) = h(f(x))) \Leftrightarrow (\forall y \in F, g(y) = h(y)) \Leftrightarrow g = h$
2. $(\cdot) \Rightarrow \exists z \neq z' \in G, (F \ni y \mapsto z) \circ f = (F \ni y \mapsto (y \in \text{Im } f ? z : z')) \circ f \therefore (\forall y \in F, y \in \text{Im } f \vee z = z') \therefore \text{Im } f = F$
3. $\text{Inj } f \Rightarrow \forall z \in G, \forall h \in G^E, h = (F \ni y \mapsto (y \in \text{Im } f ? h(f^{-1}(y)) : z)) \circ f$
4. $\forall z \neq z' \in G, \forall x \in E, \forall g \in G^F, g \circ f = (E \ni y \mapsto (y = x ? z : z')) \Rightarrow \forall y \in E, f(y) = f(x) \Rightarrow g(f(y)) = g(f(x)) = z = y = x.$

5. and 6. are particular cases of 3. ■

The case $G = V_2$ gives the following properties of $h = (\wp(F) \ni B \mapsto f_*(B))$:

$$\begin{aligned} \text{Im } f = F &\Leftrightarrow \text{Inj } h \\ \text{Inj } f &\Leftrightarrow \text{Im } h = \wp(E) \end{aligned}$$

Moreover $\{f[A] \mid A \subset E\} = \wp(\text{Im } f)$ because $\forall B \subset \text{Im } f, f[f_*(B)] = B$.

Proposition. For any sets E, F and any function g with domain F ,

1. $\text{Inj } g \Rightarrow \text{Inj}(F^E \ni f \mapsto g \circ f)$
2. $(\text{Inj}(F^E \ni f \mapsto g \circ f) \wedge E \neq \emptyset) \Rightarrow \text{Inj } g$
3. $\forall_{\text{Set}} G, (\{g \circ f \mid f \in F^E\} = G^E \wedge E \neq \emptyset) \Rightarrow \text{Im } g = G$.

Proofs:

1. $\forall f, f' \in F^E, (\text{Inj } g \wedge \forall x \in E, g(f(x)) = g(f'(x))) \Rightarrow (\forall x \in E, f(x) = f'(x))$.
2. $\forall y, z \in F, g(y) = g(z) \Rightarrow g \circ (E \ni x \mapsto y) = g \circ (E \ni x \mapsto z) \Rightarrow (\forall x \in E, y = z) \Rightarrow (E \neq \emptyset \Rightarrow y = z)$.
3. (left as an exercise to the reader). ■

Axiom of choice over a set (AC_X)

The *axiom of choice* (AC) is a statement expressible in any set theory with powerset, which «feels true» and is convenient to accept as an axiom; but we will actually not need it in the rest of this work.

It has several equivalent formulations (in clear, it is the common name of these equivalent statements). The main ones fit the common format $(\forall_{\text{Set}} X, \text{AC}_X)$ where AC_X means any of the following equivalent formulas with the free variable X in the class of sets:

- $(\text{AC1}_X) \forall_{\text{Set}} E, \forall R \subset X \times E, (\forall x \in X, \exists y \in E, xRy) \Rightarrow (\exists f \in E^X, \forall x \in X, xRf(x))$.
- $(\text{AC2}_X) \forall A : X \rightarrow \text{Set}, (\forall x \in X, A_x \neq \emptyset) \Rightarrow \prod A \neq \emptyset$
(The product of any family of nonempty sets indexed by X is nonempty)
- $(\text{AC3}_X) \forall_{\text{Fnc}} g, X \subset \text{Im } g \Rightarrow \exists f \in (\text{Dom } g)^X, g \circ f = \text{Id}_X$
- $(\text{AC4}_X) \forall_{\text{Gr}} R, X \subset \text{Dom } R \Rightarrow \exists f \in (\text{Im } R)^X, \text{Gr } f \subset R$.

(AC3_X) and (AC4_X) are more often used in cases of equality (respectively $X = \text{Im } g$ and $X = \text{Dom } R$) which are equivalent to the general cases, as the proofs of equivalence work the same :

- $(\text{AC1}_X) \Leftrightarrow (\text{AC2}_X)$ by $R = \prod A \wedge E = \text{Im } R$;
- $(\text{AC2}_X) \Leftrightarrow (\text{AC3}_X)$ by $A = (g \bullet (x))_{x \in X}$;
- $(\text{AC3}_X) \Leftrightarrow (\text{AC4}_X) : X \subset \pi_0[R] \Rightarrow \exists (h \times f) \in R^X, h = \text{Id}_X \therefore \text{Gr } f = \text{Im } (h \times f) \subset R$;
- $(\text{AC4}_X) \Leftrightarrow (\text{AC1}_X)$ obvious. ■

Other easy deductions are

- $(\text{AC2}_X) \Leftrightarrow (\text{AC1}_X)$ by $A_x = \vec{R}(x)$;
- $(\text{AC4}_X) \Leftrightarrow (\text{AC3}_X)$ by $R = {}^t\text{Gr } g$.

Dependencies between diverse AC_X

A big achievement of mathematical logic was to prove the independence of AC from the rest of set theory: even the [very strong ZF system](#), if consistent, can neither prove nor refute it, as each universe where AC is true includes another one where it is false, and vice

versa. This will be further commented in 5.5 (without the details of the proofs which are much too difficult).

The diverse ways in which AC may fail in different universes, can be partly sorted out by specifying for which X does AC_X hold. The different AC_X depend on each other as follows:

- $\forall_{\text{Set}} X, Y, (AC_X \wedge \exists f, f: Y \rightarrow X) \Rightarrow AC_Y$
- $\forall_{\text{Set}} X, Y, (AC_X \wedge AC_Y) \Rightarrow (AC_{X \cup Y} \wedge AC_{X \times Y})$
- $(AC_I \wedge \forall i \in I, AC_{A_i}) \Rightarrow AC_{\coprod_{i \in I} A_i}$

The proofs of these implications are easy and left as an exercise to the reader.

Finite choice theorem. If X is finite then AC_X .

This will be easily deducible from the case of singletons $AC_{\{x\}}$ (which is trivial) and the preservation by binary union $(AC_X \wedge AC_Y) \Rightarrow AC_{X \cup Y}$ once finiteness will be formalized (4.6). Now let us just suggest a schema of proofs, namely one for each number of elements of X . For example the case of X with 3 elements comes down to $(AC2_{V_3})$ whose proof can be written

$$\forall_{\text{Set}} X, Y, Z, (X \neq \emptyset \wedge Y \neq \emptyset \wedge Z \neq \emptyset) \Rightarrow (\exists x \in X, \exists y \in Y, \exists z \in Z, (x, y, z) \in X \times Y \times Z) \Rightarrow X \times Y \times Z \neq \emptyset. \blacksquare$$

More statements simply equivalent to AC

Theorem. The following statements are equivalent to the axiom of choice:

(AC5) Any set E of nonempty sets has a choice function: $\emptyset \notin E \subset \text{Set} \Rightarrow \prod \text{Id}_E \neq \emptyset$.

(AC6) For any partition P of a set E , $\exists K \subset E, \forall A \in P, \exists!: K \cap A$.

(AC7) For any sets E, F, G and any $g: F \rightarrow G$, $\{g \circ f \mid f \in F^E\} = G^E$.

Proofs:

(AC2) \Rightarrow (AC5) is obvious ;

(AC5) \Rightarrow (AC6) : $(x \in \prod \text{Id}_P \wedge K = \text{Im } x) \Rightarrow (K \subset E \wedge \forall A \in P, x_A \in K \cap A \wedge \forall B \in P, x_B \in A \Rightarrow A \cap B \neq \emptyset \Rightarrow A = B)$

(AC6) \Rightarrow (AC3) : $\forall_{\text{Fnc}} g, X \subset \text{Im } g \Rightarrow \exists K \subset \text{Dom } g, (\forall A \in \text{Im}(g), \exists!: K \cap A) \therefore \exists f \in (\text{Dom } g)^X,$
 $(\forall x \in X, \{f(x)\} = K \cap g \cdot (x)) \therefore g \circ f = \text{Id}_X.$

(AC1 $_E$) \Rightarrow (AC7) : $\forall h \in G^E, (\forall x \in E, \exists y \in F, g(y) = h(x)) \Rightarrow (\exists f \in F^E, \forall x \in E, g(f(x)) = h(x))$

(AC3 $_G$) \Rightarrow (AC7) : $\exists j \in F^G, g \circ j = \text{Id}_G \therefore \forall h \in G^E, j \circ h \in F^E \wedge g \circ j \circ h = h.$

(AC7) \Rightarrow (AC3) : $\forall_{\text{Set}} X, \forall_{\text{Fnc}} g, X = \text{Im } g \Rightarrow \text{Id}_X \in X^X = \{g \circ f \mid f \in (\text{Dom } g)^X\}. \blacksquare$

Note: (AC5) \Rightarrow (AC2) is also easy : $\emptyset \notin \text{Im } A \subset \text{Set} \Rightarrow \exists f \in \prod \text{Id}_{\text{Im } A}, f \circ A \in \prod A.$

There are many more statements equivalent to AC, which will not be of any use in this work, and for which the proofs of equivalence are much harder. Interested readers can find about them elsewhere.

2.A. Time in set theory

Standard universes

Given two [universes](#) U_1, U_2 , with $U_1 \subset U_2$, we shall define 3 distinct versions for the concept of U_1 being *standard* in U_2 , also called a *sub-universe* of U_2 . We first informally introduced such a concept in [1.7](#) as the preservation of the roles of sets and functions; then in [1.B](#) with the other meaning (STF) of standardness of \mathbb{N} , i.e. the interpretation of finiteness (applicable to any universe as mentioned in [1.D](#)).

The general idea is that the structures of U_1 (the considered interpretation of symbols of set theory in U_1 to make it a universe) must be the restrictions of those used for U_2 . But it depends which structures we mean, and how we define their restrictions. From the property of preservation by restriction from U_2 to U_1 we shall specify for each version, more preservation properties on other symbols implicitly follow, thanks to the axioms of set theory in U_2 . Any structure defined using only symbols with preserved meaning, is also preserved. The structures not preserved are anyway determined in U_1 as defined from those of U_2 with parameter U_1 (as a set or a unary predicate). This is why U_1 only needs to be specified as a meta-subset or class in U_2 , keeping its structures implicit.

Let us focus on specifying what each version of standardness requires on sets (to which the diversity of versions is reducible, as the conditions on functions follow in the way essentially given by the fate of their graphs as sets, after naturally defining standardness for ordered pairs). The first structure for them is their class Set . The condition for this is that $\text{Set}_1 = \text{Set}_2 \cap U_1$. Yet, arguing that Set is not really a structure but only a source of definiteness classes for effective structures, this condition may be weakened as $\text{Set}_1 \subset \text{Set}_2 \cap U_1$ without essentially affecting the main 3 degrees of standardness described below.

The weak version (coming when seeing the role of sets as given by \in) is

$$(\text{ST}) : \in_1 = \in_2 \cap (U_1 \times \text{Set}_1).$$

Actually, this condition will suffice to imply (STF) once adopted the axiom of foundation for U_2 and further reasonable conditions which will be detailed later. Otherwise we need to add it as another requirement in the concept of standardness, in the form $\mathbb{N}_1 = \mathbb{N}_2$ (if the axiom of infinity is adopted in both) or anyway in terms of the class Fin of finite sets as $\text{Fin}_1 \subset \text{Fin}_2$ which is equivalent to $\text{Fin}_1 = \text{Fin}_2 \cap \text{Set}_1$ (for reasons a bit subtle).

The intermediate version (mentioned in 1.7: every set coincides with the meta-set of the same elements) is

$$(\text{ST}') : \in_1 = \in_2 \cap (U_2 \times \text{Set}_1) \subset U_1 \times \text{Set}_1$$

i.e. every set in U_1 is interpreted by U_2 as included in U_1 . This implies the preservation of bounded quantifiers (if the sub-formula keeps its values, namely does not involve \wp) and of any operator giving a set determined by the data of its elements, namely the operators given by the set generation principle. Special constructions exist (such as those of [internal set theory](#)) of sub-universes not fitting this condition and still preserving bounded quantifiers on formulas with parameters in U_1 ; but the equivalence holds with the faithfulness of bounded quantifiers on formulas with parameters in U_2 (as we shall see when justifying the set generation principle in 2.B).

The strongest version of standardness also requires the preservation of [the powerset symbol \$\wp\$](#) (2.7):

$$(\text{STP}) : \wp_1 = \wp_2|_{\text{Set}_1}$$

This additional condition can also be written without explicit use of \wp as

$$(\text{ST}'') : (\text{ST}) \wedge \forall E \in \text{Set}_1, \forall F \in \text{Set}_2, F \subset_2 E \Rightarrow F \in \text{Set}_1$$

These conditions are related by $(\text{ST}'') \Leftrightarrow ((\text{ST}') \wedge (\text{STP})) \Rightarrow (\text{STF})$.

Indeed, (STF) comes by the fact [finiteness](#) is definable from \wp (independently of any additional assumption which was needed with the mere (ST)) ; this can also be seen considering that arithmetic is specified as a [second-order](#) theory, which can be faithfully

interpreted by set theory with \wp .

The independence of the axiom of choice (further [commented in 5.5](#)) and other statements involving \wp , can be roughly explained by the possibility to construct universes which do not preserve the essential meaning of \wp . Following convenience, one such construction happens to fit (ST'); others may fit (ST) \wedge (STP) but not (ST') while (ST')-standard copies of their results also exist.

Ideally standard universes

For any 3 universes $U_1 \subset U_2 \subset U_3$ where U_2 is standard in U_3 ,

$$(U_1 \text{ is standard in } U_2) \Leftrightarrow (U_1 \text{ is standard in } U_3).$$

Thus the idea to consider the standardness of U_1 as an absolute property, independent of the external universe in which it is described... provided that this external universe is itself standard. This does not define standardness as an absolute concept (which cannot be defined anyway), but suggests that such a concept may make ideal sense.

A sub-universe U_1 of U_2 will be qualified as *small* in U_2 if it is known as a set there. This can be written $U_1 \in \text{Set}_2$ if U_2 is also itself (ST')-standard, otherwise $\exists X \in \text{Set}_2, U_1 = \overset{\leftarrow}{\epsilon}_2(X)$.

To complete the concept of ideally standard universe, let us postulate that every two standard universes are small sub-universes of a third one.

This ideal is quite reasonable for (ST) and (ST'); but for (ST'') it is somewhat more daring. As a first reason, it can be philosophically argued that any universe (standard or not) can be taken as a small sub-universe of another in the sense (ST'), but not always in the sense (ST'') because seeing it as a set may provide means to define more subsets of given sets. In particular, any (ST'')-standard universe being a small (ST'') sub-universe of another, implies its fulfillment of the [specification schema](#).

The realistic view of set theory

From now on we shall assume a fixed version of "standardness" meant as either (ST') with a set theory without powerset, or (ST'') with a set theory with powerset. In either case, all bounded expressions keep the same values from a standard universe to another, thus called their standard values.

The above concepts provide a [realistic meaning](#) for set theory :

- Generic theories, seen axiomatically, are meant to be interpreted in a fixed model containing all values of variables and expressions;
- But set theory aims to "locally" interpret in the standard way any expression with fixed values of its free variables, independently of the choice of a standard universe, containing these values, where this interpretation may be processed.

In the next section (2.B) the meaningfulness of the (ST'') ideal will appear philosophically equivalent to the realism of admitting \wp in set theory.

Let us formalize some aspects of the plurality of standard universes which set theory aims to describe in its realistic view.

Standard multiverses (expansion of the set theoretical universe)

Let us call *quasi-standard multiverse* (resp. *standard multiverse*) any collection (range) of universes (resp. standard universes), where any 2 of them are small sub-universes of a third one. Let us require it to also contain all sub-universes of its members, a quality easy to fulfill, taking a multiverse where it did not hold and adding all these sub-universes to it.

We shall say that the set theoretical universe *expands* when it ranges over a standard multiverse.

For any quasi-standard multiverse M , its [union](#) $\mathbf{U} = \bigcup M$ forms another universe containing all members of M as small sub-universes. Then, for any expression with any values of its free variables in \mathbf{U} there exists some $U \in M$ containing these values and thus able to interpret the expression; this interpretation being independent of the chosen U , defines that of \mathbf{U} . Therefore,

$$(M \text{ is a standard multiverse}) \Leftrightarrow (\text{the universe } \mathbf{U} \text{ is standard}).$$

But \mathbf{U} cannot belong to M , because M has no greatest element. So, no single standard multiverse can ever contain all possible standard universes.

For a (ST^{''})-quasi-standard multiverse, keeping in \mathbf{U} the same powerset of a given set as in member universes, may lead to the failure of the specification schema in \mathbf{U} . Here comes another side of the postulate attached to the ideal (ST^{''})-standardness : that with (ST^{''})-standard multiverses, this failure never happens.

Can a set contain itself ?

Let us call a set *reflexive* if it contains itself.

The [proof of Russell's paradox](#) [1.8], shows that the class of non-reflexive sets ($\text{Set}(x) \wedge x \notin x$) cannot be a set (as it is not included in any set). So, many sets are non-reflexive ; but can reflexive sets exist ? this is undecidable; here is why.

From any universe U with a nonempty class X of reflexive sets, another universe containing none can be made by 2 different methods

- Take $U \setminus X$ as class of objects, and turn into urelements all sets E which contained a reflexive set ($\exists x \in E, x \in x$), and all functions whose domain or image was such an E . (These sets cannot be rebuilt from the data of their elements, since each one has an element removed from the universe, namely itself ; any class of urelements can also be eliminated in the same way, turning other objects into urelements).
- Progressively rebuild a universe from scratch: this avoids all reflexive sets. Namely, from a given universe, first accept its urelements, then also accept, progressively along an endless "time", every set only made of previously accepted objects, and every function only involving such objects. Each set so accepted at some time, is formed as a collection of previously accepted objects. At the first time a set is so accepted as a collection of previously available objects, it could not be available yet as a possible element, thus it cannot be reflexive.
As the reflexivity of a set is independent of context, a union of universes each devoid of reflexive sets, cannot contain one either.

On the other hand, universes with reflexive sets can be produced as follows:

Riddle. What is the difference between

- a universe with an urelement x and a non-reflexive set y such that $x \in y$,
- and a universe with a reflexive set x and an urelement y such that $y \notin x$?

► **Answer** (click to show) :

The role of the set containing x but not y , played by y in the former universe, is played by x in the latter.

The axiom of foundation, introduced in [5.3 \(well-foundedness\)](#), will imply the absence of reflexive sets as a particular case ; its undecidability can be proven in ways essentially similar to the above arguments. But for most purposes, this axiom is as useless as the sets it excludes.

2.B. Interpretation of classes

The relative sense of open quantifiers

As the set theoretical universe [expands](#), the values of [decidable statements](#) (under the given [axioms](#)) must of course stay constant. However, any undecidable statement, as well as any other formula containing open quantifiers and given with fixed values of its free variables (let us call these "open formulas"), may change value from one universe to another.

But if any value of an open formula can only be "the current one" relative to the choice of this universe, the question whether it is [constant](#) or variable, that is how things go «elsewhere», is relative to a choice of multiverse. But the [realistic view](#) of set theory means to neither assume a fixed universe nor a fixed multiverse. Thus, the natural way for it is to abstain from giving any value to open formulas (beyond decidable statements with known proof or refutation). Indeed, tries to do better would be rather hopeless due to how messy the situation is. Let us review its diverse aspects.

In a multiverse of «all universes» (standard or not) fitting a given consistent axiomatic set theory, the situation was explained in Part 1 :

- The variability of value of a statement there, is equivalent to its undecidability (completeness theorem);
- However, this undecidability, when true, cannot anyway be proven by the same axiomatic set theory (incompleteness theorem).

The realistic view referring to standard universes, leads to another messy situation, first by lack of formal definition for the ideal concept of standardness.

Trying to materialize the concept of "real indefiniteness" of statements as their variability in a standard multiverse M , the behavior of this M is determined by its union \mathbf{U} , as M is the range of all small sub-universes of \mathbf{U} ; this allows to express relevant questions as statements in \mathbf{U} .

For example, the variability in M of an [existential](#) statement ($\exists x, A(x)$ where A is a bounded formula) means there exists in M a universe U where it is false ($\forall x \in U, \neg A(x)$), and another universe U' where it is true ($\exists x \in U', A(x)$). But then it is also true in any standard universe containing both U and U' , among which some other members of M , and \mathbf{U} itself. Such a statement may be called «ultimately true».

But since some statements may vary between acceptable standard universes U , some of these may still vary between the \mathbf{U} of acceptable standard multiverses. Indeed, \mathbf{U} is just another universe which can be axiomatically described, but only incompletely, for both formal reasons (the incompleteness theorem) and realistic ones (there is no ultimate standard multiverse). Just more statements can be made decidable for \mathbf{U} than for U by describing \mathbf{U} with a stronger theory \mathbf{T} than the theory T given to describe U . Such a stronger \mathbf{T} can already be obtained by formalizing this description of \mathbf{U} as the union of a quasi-standard multiverse whose members satisfy T (while the standardness of \mathbf{U} escapes formalization).

Behind the undecidability of a statement (existential or other), multiple situations may occur. It might (I guess) be always true in one quasi-standard multiverse, and always false in another, while, depending on the statement, either of them (but not both) may be standard. It may vary or not vary between standard universes, but this variability, itself expressible as a statement in the union of a standard multiverse, may be itself undecidable, and eventually vary between standard multiverses.

The undecidability of some statements such as AC, usually reflects the variability, which may also happen among (ST')-standard universes, of the values of the open quantifiers used in these statements when [re-written without \$\wp\$](#) . In particular, a quantifier $\forall x \in \wp(E)$ appears as an open quantifier on the class of subsets of E , while other uses of \wp either also

involve this once translated, or cannot even be translated when the powerset axiom does not hold.

The indefiniteness of classes

Unlike objects which can be compared in formulas by the = symbol, the meta-relation of equality between classes is as indefinite as the open \forall since both concepts are definable from each other :

- The equality of classes $A = B$ would be defined by $\forall x, A(x) \Leftrightarrow B(x)$;
- The statement $(\forall x, A(x))$ means $A = 1$ (the universe, class of all objects).

This indefiniteness can be understood by remembering that, despite being usable as ranges, classes are only meta-objects, basically given as predicates, i.e. syntactically (as a formula with parameters). Like with open quantifiers, this leaves us with both concepts of *provable equality* (or proven equality), and *provable inequality*, according to the status (provable or refutable) of this statement of "equality" of classes $(\forall x, A(x) \Leftrightarrow B(x))$.

Each universe U interprets each class C as a meta-set of objects $P = \{x \in U \mid C(x)\}$, and sees it as a set when $P \in U$. This condition is expressible by set theory in U as a statement $S(C)$, written equivalently as

$$\begin{aligned} \exists P, \forall x, C(x) \Leftrightarrow x \in P \\ \exists E, \forall x, C(x) \Rightarrow x \in E \end{aligned}$$

The equivalence is because P can be defined from such an E by $P = \{x \in E \mid C(x)\}$. Having two open quantifiers, it is "more indefinite" than the equality between classes.

Classes and sets in expanding universes

Aside this formulation $S(C)$ in a fixed universe, let us analyze this distinction of sets among classes (and thus what makes other classes indefinite), from the perspective of a quasi-standard multiverse. There, this concept has 3 possible versions, ordered only approximately by implications which "often work" but have exceptions.

- The "weak" version is to claim $S(C)$ true in all members of this multiverse. That means C always coincides with some set P , but this may be a different P from one universe to another.
- The "strong" version requires P to stay constant. This can be seen as the meaning behind the formal choice of naming P by a symbol of set theory, as was discussed with the powerset (2.7). But just the constancy of $P = \{x \in U \mid C(x)\}$ when U varies, needs not imply the above weak version, as there can exist some universe U (among the smallest of this multiverse) such that $P \notin U$. This means P is arising to existence with U , and will anyway exist as a set in all the "future" (larger) universes which see U as a set. So, a class C is not considered a set in this "strong" sense, if it remains able to contain «unknown» or «not yet existing» objects (in some future universe), that would only belong to some larger future value of P , which can thus vary with the growth of U .
- The middle version (implied by the strong one) is to interpret $S(C)$ in the union \mathbf{U} of the multiverse. Deciphered there, it says that, while U expands, «there exists a time from which P stays constant» (and thus after which it will exist as a set). So, this just differs from the strong version by ignoring any "past" variations (which could occur during some "beginning" of the expansion), to focus on the ultimate behavior (among the latest, i.e. largest universes, somehow "closest" to the \mathbf{U} where $S(C)$ is interpreted).

But the ultimate realistic meaning of set theory is to refer to the range of "all standard universes", which differs from the concept of standard multiverse in that this range is not

a set but a class. So, the concept of union cannot apply to this range, and yet its impossible result would form the ideal interpretation of set theory. This ideal can be well approximated by interpretations of set theory by either fixed universes or quasi-standard multiverses (whose existence is ensured by the completeness theorem). Both perspectives, in terms of a fixed vs. a variable universe, alternatively transcend each other endlessly along its expansion.

This explains how any intended set theory can be formalized by mere first-order axioms: if any intended property of the universe was only expressible by a second-order statement, or if anyhow its expression involved external objects (regarding this universe as a set), then it could be re-expressed by moving the framework, as the stronger first-order axiom of existence of a sub-universe of this kind, and why not also endlessly many of them, forming a standard multiverse (stating that every object is contained in such a sub-universe).

Justifying the [set generation principle](#)

Let B be a quantifier defined by a bounded formula whose parameters take values in a sub-universe U of any other universe we might work in. Let E be the range of all values taken by the argument y of A when interpreting $(By, A(y))$. It is independent of A and included in U insofar as A is absent from any expression of y in the defining formula. Typical exceptions would be when B involves a formula of the form $A(t(\{x \in F \mid A(x)\}))$ for some term t , in which case defining E as the set of *possible* values of y would require to assume \varnothing and the (ST^U)-standardness of U . Without these assumptions, such exceptional formulas need to be excluded from the definitions of B for the justifications to work. Anyway they are excluded by understanding, as explained in 2.A-2.B, the set-builder as the notation for the operator symbols from an infinite list, one for each formula using the language of set theory to which A does not belong. (The set-builder is needed to put a formula inside a term, unless the conditional operator is accepted as primitive, which is acceptable, but then, function definers using A by the conditional operator are excluded for the same reason)

Let $C(x)$ defined as $(By, y = x)$. By the hypothesis of the set generation principle for the B which was written Q^* , we have a proof of $(B \Leftrightarrow \exists C)$. This implies $\neg(B, 0)$, and, coming by second-order universal introduction, remains valid in any universe where it is interpreted.

For any x , the value $C(x)$ of B on the predicate $(y \mapsto (y = x))$, can only differ (be true) from its (false) value on $(y \mapsto 0)$, if both predicates differ inside E , i.e. if x belongs to E :

$$C(x) \Leftrightarrow ((By, y = x) \Leftrightarrow (By, 0)) \Rightarrow (\exists y \in E, y = x \Leftrightarrow 0) \Leftrightarrow x \in E$$

This inclusion of C in E shows it is essentially a set:

- C is a set in the above "strong" sense ($C \subset U$) ;
- Moreover E is [a set in the sense of first-order logic](#) (1.D): the formula of B only has definite, fixed means (variables bound to given sets, fixed parameters) to provide its elements. ■

For any class satisfying the condition of the set generation principle (being indirectly as usable as a set in the role of domain of quantifiers), is it also indirectly as usable as a set in the role of domains of functions (before using this principle) ? Namely, is there for each such class a fixed formalization (bounded formulas with limited complexity) playing the roles of the definer and the evaluator for functions having this class as domains ? The answer would be yes, but we shall not develop the justifications here.

Concrete examples

A set: Is there any [dodo](#) left on Mauritius ? As this island is well known and regularly

visited since their supposed disappearance, no surviving dodos could still have gone unnoticed, wherever they may hide. Having not found any, we can conclude there are none. This question, expressed by a bounded quantifier, has an effective sense and an observable answer.

A set resembling a class: Bertrand Russell raised this argument about theology: «If I were to suggest that between the Earth and Mars there is a china teapot revolving about the sun..., nobody would be able to disprove my assertion [as] the teapot is too small to be revealed even by our most powerful telescopes. But if I were to go on to say that, since my assertion cannot be disproved, it is intolerable presumption on the part of human reason to doubt it, I should rightly be thought to be talking nonsense.» The question is clear, but on a too large space, making the answer practically inaccessible. (An 8m telescope has a resolution power of 0.1 arcsec, that is 200m on the moon surface)

A class: the extended statement, «there is a teapot orbiting some star in the universe» loses all meaning: not only the size of the universe is unknown, but [Relativity theory](#) sees the remote events from which we did not receive light yet, as not having really happened for us yet either.

A meta-object: how could God «exist», if He [is a meta-object](#), while «existence» can only qualify objects? Did apologists properly conceive their own thesis on God's «existence»? But what are the *objects* of their faith and their worship? Each monotheism rightly accuses each other of only worshipping objects (sin of idolatry): books, stories, beliefs, teachings, ideas, attitudes, feelings, places, events, miracles, healings, errors, sufferings, diseases, accidents, natural disasters (declared God's Will), hardly more subtle than old statues, not seriously checking (by fear of God) any hints of their supposed divinity.

A universal event: the redemptive sacrifice of the Son of God. Whether it would have been theologically equivalent for it to have taken place not on Earth but in another galaxy or in God's plans for the Earth in year 3,456, remains unclear.

Another set reduced to a class... The class F of girls remains incompletely represented by sets: the set of those present at that place and day, those using this dating site and whose parameters meet such and such criteria, etc. Consider the predicates B of beauty in my taste, and C of suitability of a relationship with me. When I try to explain that «I can hardly find any pretty girl in my taste (and they are often unavailable anyway)», i.e.

$$(\forall_F x, C(x) \Rightarrow B(x)) \wedge \{x \in F \mid B(x)\} \approx \emptyset,$$

the common reaction is: «Do you think that beauty is the only thing that matters?», i.e.

$$\text{What, } (\forall x \in F, C(x) \Leftrightarrow B(x)) \text{ ???}$$

then «If you find a pretty girl but stupid or with bad character, what will you do?».

Formally : $(\exists x \in F, B(x) \neq C(x) !!!)$. And to conclude with a claim of pure goodness: «I am sure you will find !», that is $(\exists \text{ plenty of } x \in F, C(x))$. Not forgetting the necessary condition to achieve this: «You must change your way of thinking».

... by the absence of God...: F would have immediately turned into a set by the existence of anybody on Earth able to receive a message from God, as He would obviously have used this chance to make him email me the address my future wife (or the other way round).

... and of any substitute: a free, open and efficient online dating system, as would be included in my project [trust-forum.net](#), could give the same result. But this requires finding programmers willing to implement it. But the class of programmers is not a set either, especially as the purpose of the project would conflict with the religious moral priority of saving God's job from competitors so as to preserve His salary of praise.

Another concrete example of a small proper class is given by [Laureano Luna](#).

2.C. Concepts of truth in mathematics

Let us recall and expand on the different concepts of «truth» we saw for mathematical statements, from the simplest to the most subtle.

The most basic was the **relative truth** (further analyzed in [1.B](#)), that is the Boolean value of any statement expressed in any given theory, depending on the system (model of the theory) where it is interpreted; this system is supposedly fixed like an implicit free variable. More generally, any formula takes a Boolean value for fixed values of its free variables in that system.

Then comes **provability** in any given first-order axiomatic theory ; this coincides with the quality of being relatively true in all its models.

Finally, come the truths in two realistic theories : arithmetic and set theory. Let us review them in more details.

Arithmetic truths

As [explained in 1.C](#), the concept of provability in any explicit first-order theory, and more generally any effective "provability" concept (predicate) over the class of statements of any conceivable theory, must be a **FOT**-existential predicate, equivalent (through some encoding) with some existential predicate of arithmetic. But we need to refer to **the realistic truth of arithmetic** to interpret these. This is an ideal but clear concept (independent of any ontological hypothesis about infinities), meaning the class of properties of "the standard model \mathbb{N} of arithmetic" (to say in short, as all standard models of arithmetic are identical copies of each other).

Indeed all standard objects of a FOT being finite can be accepted as real (no matter how big), while their distinction with non-standard objects (pseudo-finite ones) is also real despite being not formalizable.

The meaningfulness of statements can be understood first for bounded formulas, then those with 1 open quantifier, such as provability or consistency statements, whose objects are mere finite systems (proofs). And so on for each additional open quantifier over a previously accepted formula, even if their meaningfulness becomes more and more subtle.

Yet the incompleteness theorem showed that this realistic truth predicate of arithmetic is not itself an existential predicate. More precisely, even the negation of the provability predicate of any explicit theory able to express arithmetic, cannot be itself an existential predicate.

Still we need partial solutions: existential classes of statements of arithmetic with both qualities

- FOT-sound (included in the class of realistic truths);
- Large (compared by inclusion with other such existential classes).

The natural way to progress in the endless and non-algorithmic quest for larger and larger such classes without breaking FOT-soundness, consists in searching for stronger and stronger axiomatic set theories. Let us comment this further.

More concepts of strength

Any axiomatic set theory (or other foundational theory) T is reflected by 2 classes of arithmetical statements :

1. The class of its arithmetical theorems ;
2. The statement of its consistency.

Both should not be confused :

- 1. cannot contain 2. (unless 2. is false), though 2. must be true if 1. is valid. This is because 1. is only made of the theorems of T with a standard proof, while 2. means to forbid the current universe from containing even a non-standard proof (with pseudo-finite length) from T that would lead to contradiction.
- So, 2. means the existence, in the current universe, of some model of T but with no requirement of standardness ; while 1. means the current universe agrees with T on arithmetic, and thus can only be philosophically justified by the idea that T is valid in the sense of having a FOT-standard model.

They lead to two other conceptions (definitions) of a "strength" preorder between foundational theories, most often equivalent to our first one of [1.A](#) (possible cases of non-equivalence will not be considered here) :

1. The inclusion order between their classes of arithmetical theorems ;
2. The deducibility (provable implication) between consistency statements: T' is stronger than T if the consistency of T is deducible from the consistency of T' ; then T' is "strictly stronger" than T if the consistency of T is a theorem of T' .

(1. would be related to the implication order between statements of FOT-soundness, except that the expression of such statements requires a framework strong enough to express arithmetical truth, such as MT).

If T' can prove the existence of a standard model of T then T' is strictly stronger than T in both ways.

Set theory from realism to axiomatization

For an axiomatic set theory to give a class of arithmetical theorems both FOT-sound and large, it needs to be "not wrong and yet very good" :

- Sound = keeps accepting some standard universes (for FOT-soundness);
- Strong = only accepts "large" universes (truly large in the standard case, otherwise anyway looking large by the description of their hierarchy of sub-universes), to contribute to a large class of theorems.

Among sound, well-described set theories (or other foundational theories) there can be no strongest one : from any of them T we can get stronger ones, at least arithmetically, in the following ways roughly ordered by increasing power (where T_0 is some possibly weaker theory but which fits the below "open" quality):

- T + the consistency of T ;
- T_0 + the existence of a standard universe of T ;
- T_0 + the existence of infinitely many standard universes of T forming a standard multiverse (actually only ensured to be quasi-standard).

And these are only the first of an open-ended range of possible methods with growing efficiency in their strengthening effect with respect to the added complexity of description. It actually turns out that the Replacement schema, used in ZF set theory, amounts to the use of a [much more powerful strengthening method](#) than these.

Arguments to justify any strong axiomatic set theory like this as a valid foundation of mathematics, must remain somewhat philosophical, and thus assessed in intuitive, not completely formalizable ways, since their point is to do better than any formal proof or other predefined algorithm: no fixed formal method can always witness for any strong theory when it is consistent, even less when it is FOT-sound (by truth undefinability), or sound (the idea that some given axiom which excludes one size of standard universe still accepts some larger ones).

The usefulness of strong axiomatic set theories for proving large classes of arithmetical truths (insofar as philosophical disputes on the ontological status of the ideals motivating these theories do not undermine the role of these ideals as reasons for the confidence in the FOT-soundness), can then be read as an indispensability argument for the reality of the universes so described, beyond the infinity of \mathbb{N} .

Indeed, it makes these theories "valuable", while their consistency together with the mere existence of \mathbb{N} , ensure the existence of models. Only non-standard models are so ensured to exist, but these work similarly to the standard ones which were ideally meant anyway. Hence our last concept of truth, which is the **truth of set theoretical statements**.

Axioms compatibility condition

Given several axiomatic set theories with both qualities (sound and strong), consider the theory obtained as their union (the union of their classes of axioms if they have the same symbols, for example if all only use \in , like ZF). It will be stronger than each, but is it still sound? Let us qualify a collection of set theories (or their axioms) as *compatible* if their union (conjunction) remains sound.

So, beyond the soundness and strength of axiomatic set theories, we need an additional quality requirement for their axioms, so that, without limiting the strength of sound set theories which can be so written, all such theories will be compatible. Actually this problem has a natural solution. This needed additional quality is another ideal concept, not completely formalizable, so let us first express it intuitively :

- An axiom is *open* if satisfied by any *open universe*, i.e. where eternity is a very long time **especially towards the end**.

This is meant as more specific than the mere "eternity is a very long time" which rephrases the quality "large" for universes, and as deepening our previous concept of being [open](#). Let us explain it in several steps.

First, no axiom should put any limit on the size of the universe, for compatibility with set theories stronger than this size (for example, a sound theory T accepting only one standard universe would be incompatible with the statement of existence of a standard universe of T).

Then, the remaining risk for several statements to be incompatible, is if the values of at least 2 of them endlessly vary (alternating truth and falsity) as the universe expands, in such a way that no standard universe "larger than some size" can fit all (their conjunction would limit the size of the universe). This is resolved by only accepting "open" axioms, which will all agree on "open" standard universes. Let us explain how.

Consider such a variable statement : written in [prenex form](#) it must be using both kinds of open quantifiers. Let us analyze the case of statements with only 2 open quantifiers: $\exists x, \forall y, A(x,y)$, among which the statements [S\(C\)](#) that some class is a set. (From this, the case $\forall x, \exists y$ is deduced by negation, while cases with more open quantifiers would be trickier and will not be discussed here). For any standard multiverse where it indeed endlessly varies, such a statement turns out to be false in its union. So, taking such a statement as an axiom would be inappropriate for this multiverse. Yet, what really matters is the behavior of the class of all standard universes (there is no systematic way to determine this, but it is our intended ideal). Now the point of a statement being "open", is that it reflects the truth in the union of a multiverse which behaves like the intended class of all standard universes.

Such "open" universes, with the same properties as the union of a standard multiverse behaving like the class of "all standard universes", can be intuitively described as "much larger than any smaller one" : there must be no limit to how bigger the universe is from any sub-universe (described by a weaker theory). This also avoids the inefficiency that could be involved in the opposite case, of an axiomatic system likely to be significantly

more complex but only slightly stronger than another one.

The quest for axiomatic set theories approaching all 3 ideally conceived qualities (strong and open but still sound) aiming to select universes with the corresponding 3 qualities (large and open but still existing among standard ones), is endless. Fortunately, rather simple set theories such as ZF, already satisfy these qualities to a high degree, describing much larger realities than usually needed. This is how a Platonic view of set theory (seeing the universe of all mathematical objects as a fixed, exhaustive reality) can work as a good approximation, though it cannot be an exact, absolute fact.

Alternative logical frameworks

The description we made of the foundations of mathematics (first-order logic and set theory), is essentially just an equivalent clarified expression of (or other way to introduce) the same version of mathematics as widely accepted. Other [logical frameworks already mentioned](#), to be developed later, are still in the "same family" of "classical mathematics". But other, more radically different frameworks (concepts of logic and/or sets), called *non-classical logic*, might be considered. Examples:

- Some logicians developed the «intuitionistic logic», which lets formulas keep a possible indefiniteness as we mentioned for open quantifiers, but treated as a modification of the pure Boolean logic (the rejection of the excluded middle, where $\neg(\neg A)$ does not imply A), without any special mention of quantifiers as the source of this uncertainty. Or it might be seen as a formal confusion between truth and provability. In this framework, $\{0\} \cup]0,1] \subset [0,1]$, without equality. I could not personally find any interest in this formalism but only heard that theoretical computer scientists found it useful.
- When studying [measure theory](#) (which mathematically defines probabilities in infinite sets), I was inspired to interpret its results as simpler statements on another concept of set, with the following intuitive property. Let x be a variable **randomly** taken in $[0,1]$, by successively flipping a coin for each of its (infinity of) binary digits. Let E be the domain of x , set of all **random** numbers in $[0,1]$. It is nonempty because such random numbers can be produced. Now another similarly random number, with the same range ($y \in E$) but produced independently of x , does not have any chance to be equal to x . Thus, $\forall x \in E, \forall y \in E, x \neq y$. This way, $x \in E$ is no more always equivalent to $\exists y \in E, x = y$.

We will ignore such alternatives in the rest of this work.