

CDM

Sets

Klaus Sutner
Carnegie Mellon University
www.cs.cmu.edu/~sutner

Battleplan

- Intuitive Set Theory
- Cartesian Products
- General Products and Sums
- Cardinality
- Frege's System
- Russell's Bomb
- Zermelo-Fränkel Set Theory

Intuitive Set Theory

A Definition of Set

Definition 1. A **set** is an arbitrary collection of objects.

In the words of Georg Cantor:

By an “aggregate” we are to understand any collection into a whole M of definite and separate objects m of our intuition or our thought. The objects are called “elements” of M . In signs we express this thus: $M = \{m\}$.

Cantor’s symbolic notation is rather old-fashioned. Nowadays one would usually write

$$M = \{ m \mid P(m) \}$$

indicating that we wish to collect all objects m that have property P into a set M .

But Why?

As it turns out, sets alone (well, plus a bit of logic) suffice as a foundation of more or less all of mathematics and computer science. Set theory provides an extremely powerful and even elegant way to organize and structure any discourse in this domain.

On the face of it, this is a huge surprise: one would suspect that sets are nowhere near powerful enough to express concepts such as natural number, prime, group, field, real number, differentiable function, probability measure, finite state machine, computable function, complexity class, and so on.

The importance of sets can be seen in the fact that Bourbaki chose to dedicate his first volume to sets:

- I Set Theory
- II Algebra
- III Topology
- IV . . .

Urelements

Of course, there is an immediate question: sets of what?

For example, in arithmetic we would like to have sets of natural numbers such as the primes. In calculus we would like to have sets of reals such as intervals. In algebra we are dealing with groups, rings, fields, polynomials, matrices and so on.

But in pure set theory none of these objects exist.

One way around this problem is to assume that we are given a collection \mathcal{U} of *urelements*. We are allowed to form subsets of \mathcal{U} , subsets of subsets, combinations with pure sets, and so on.

Different sets of urelements are appropriate for different areas of discourse. The approach is quite practical and is often used tacitly, without any mention of the underlying idea.

Pure Sets

As we will see below, urelements are superfluous in the sense that the “stuff of mathematics and computer science” can always be represented in terms of pure sets. Unfortunately, a complete definition of say, the real numbers, in terms of sets only is quite unwieldy.

One winds up with complicated, infinite sets of sets of sets of sets . . . that have at best a tenuous connection to anyone’s intuition about what a real number is. But, one can then give a rigorous proof that the reals are complete.

The real purpose of set-theoretic definitions is to provide a precise standard, a solid reference for all the more informal notions that one uses in actual practice.

Higher levels of abstraction are indispensable for any real application, but set theory provides the bedrock foundation.

Sets as Data Structures

We can think of a set as a kind of general purpose data structure – a container type.

Since we are dealing with arbitrary collections there are several natural operations:

- insert, remove, membership test, . . .
- union, intersection, difference, size, . . .

Warning: Sets are very different from sequential collections such as lists or arrays:

- Sets are not ordered, there is no first, second, . . . , last element.
- There are no multiple elements.

As a matter of fact, sets are notoriously difficult to implement; sequential containers such as lists are much easier to deal with.

Intuitive Set Theory

Standard notation for finite sets treats them very much like lists, as a sequential container:

$$S = \{a_1, a_2, \dots, a_{n-1}, a_n\} \quad \text{set formation}$$
$$\emptyset = \{\} \quad \text{empty set}$$

But note that this notation is a bit misleading in that we must have

$$\{a, b, c\} = \{b, a, a, c, a, c, b\}$$

Duplicates and order are irrelevant for sets. This is the crux in implementing sets rather than plain lists; we have to make an extra effort to avoid duplication and to ignore order (though elements in a data structure are always ordered in some way, e.g. by the order of insertion which is, of course, extraneous to the actual collection).

Membership Relation

There is only one fundamental relation between sets: *membership*.

Notation

$$x \in y \quad x \text{ is an element of } y.$$

Example 1.

$$5 \in \{2, 3, 5, 8\}$$

$$7 \notin \{2, 3, 5, 8\}$$

$$z \in [0, 1) \iff 0 \leq z < 1$$

$$x \notin \emptyset \text{ for any } x$$

As we will see in a moment, for sets even equality can be reduced to membership.

Infinite Sets

For infinite collections that have a very simple, regular structure we can still use the same notation, augmented by ellipses:

$$\text{Even} = \{\dots, -2n, \dots, -2, 0, 2, \dots, 2n, \dots\}$$

$$\text{Primes} = \{2, 3, 5, 7, \dots, 1299709, \dots\}$$

Note, though, that these definitions depend on the ability of the reader to interpret the dots. As input to a program ellipses are usually not allowed (except in very restricted circumstances such as to denote intervals of integers).

In the examples above, we presumably have the integers as urelements, or we could represented them as pure sets, see below.

Set Formation

To obtain complicated sets we can collect all objects z (all of them sets or urelements) with a certain property $P(z)$ into one set:

$$A = \{ z \mid P(z) \}$$

Very often one selects elements from some larger collection B that has already been constructed.

$$A = \{ z \in B \mid P(z) \}$$

Example 2.

$$[n] = \{ z \in \mathbb{N} \mid 1 \leq z \leq n \}$$

$$\mathbb{P} = \{ z \in \mathbb{N} \mid z \text{ is prime} \}$$

$$[0, 1) = \{ z \in \mathbb{R} \mid 0 \leq z < 1 \}$$

$$\mathbb{Q} = \{ a/b \mid a, b \in \mathbb{Z}, b > 0 \}$$

Set Operations

There are several simple operations on sets that are useful to construct new sets from given ones.

union	$A \cup B$	$= \{ x \mid x \in A \vee x \in B \}$
intersection	$A \cap B$	$= \{ x \mid x \in A \wedge x \in B \}$
difference	$A - B$	$= \{ x \mid x \in A \wedge x \notin B \}$
symmetric diff.	$A \Delta B$	$= \{ x \mid x \in A \oplus x \in B \}$

Here \oplus means exclusive or. So all this boils down to simple propositional logic.

Example 3.

$$\{1, 2, 3\} \cap \{2, 3, 4, 5\} = \{2, 3\}$$

$$\{1, 2, 3\} \cup \{2, 3, 4, 5\} = \{1, 2, 3, 4, 5\}$$

$$\{1, 2, 3\} \Delta \{2, 3, 4, 5\} = \{1, 4, 5\}$$

Basic Set Properties

- Associativity

$$x \cup (y \cup z) = (x \cup y) \cup z \text{ and}$$

$$x \cap (y \cap z) = (x \cap y) \cap z.$$

- Commutativity

$$x \cup y = y \cup x \text{ and } x \cap y = y \cap x.$$

- Distributivity

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z) \text{ and}$$

$$x \cup (y \cap z) = (x \cup y) \cap (x \cup z).$$

- Identity

$$x \cup \emptyset = x \text{ and } x \cap \mathcal{U} = x.$$

- Idempotence

$$x \cup x = x \text{ and } x \cap x = x.$$

- Absorption

$$x \cup (x \cap y) = x \text{ and } x \cap (x \cup y) = x.$$

Exercise 1. *Show that, intuitively, all these properties hold.*

Complements

In general “the complement of set x ” makes no sense: before we can remove the elements of x we have to determine what they should be removed from. In other words, we need to fix some universe \mathcal{U} and consider only $x \subseteq \mathcal{U}$:

$$x^- = \mathcal{U} - x.$$

Now assume $x, y \subseteq \mathcal{U}$ for some fixed universe \mathcal{U} . Then we have

- Domination
 $x \cup \mathcal{U} = \mathcal{U}$ and $x \cap \emptyset = \emptyset$.
- Complements
 $x \cup x^- = \mathcal{U}$ and $x \cap x^- = \emptyset$.
- Double Complement (involution):
 $x^{--} = x$.
- De Morgan's Laws:
 $(x \cup y)^- = x^- \cap y^-$ and $(x \cap y)^- = x^- \cup y^-$

Déjà Vu All Over Again?

These rules should all look eminently familiar: they are essentially the axioms for a Boolean algebra.

Of course, this is no coincidence: Fix some universe \mathcal{U} and interpret $+$ by \cup , \cdot by \cap , and \bar{x} as $x^- = \mathcal{U} - x$. We obtain a structure

$$\langle \mathfrak{P}(\mathcal{U}), \cup, \cap, ^-, \emptyset, \mathcal{U} \rangle$$

Lemma 1. *The powerset of \mathcal{U} with the operations union, intersection and complement forms a Boolean algebra.*

Inquisitive minds might wonder why true/false should behave just like all subsets of a fixed universe \mathcal{U} . We'll come up with a fairly good explanation in a while.

Unary Operations

Definition 2. Let X be a set (intended: a set of sets).

$$\bigcup X = \{z \mid \exists x (z \in x \wedge x \in X)\}$$

$$\bigcap X = \{z \mid \forall x (x \in X \rightarrow z \in x)\}$$

These definitions are a bit easier to read if we write

$$\bigcup X = \{z \mid \exists x \in X (z \in x)\}$$

$$\bigcap X = \{z \mid \forall x \in X (z \in x)\}$$

Exercise 2. Show that $\bigcup \{a, b\} = a \cup b$ and $\bigcap \{a, b\} = a \cap b$.

Example: Generating Subgroups

A typical example of the use of the general intersection operator is to guarantee the existence of a subgroup generated by some elements $A \subseteq G$, where G is some group.

In this case let

$$X = \{ H \subseteq G \mid A \subseteq H, H \text{ subgroup} \}$$

Then X is not empty since certainly $G \in X$.

But subgroups are closed under intersection, so $\bigcap X$ must be the subgroup we are after.

Note that from a certain perspective this argument is a bit circular: $\bigcap X \in X$, so we are using an object to define itself. This is called an impredicative definition.

Extensionality

Extensionality Principle

Two sets are considered to be the same iff they contain precisely the same elements.

$$x = y \iff \forall z (z \in x \leftrightarrow z \in y)$$

The importance of this principle was first recognized by Leibniz and is enshrined in his *principium identitatis indiscernibilium*: if we cannot tell two entities apart they are identical. Later G. Frege incorporated this principle in his system as the infamous Axiom V, a decision that would cause him major headaches.

It is a consequence of Extensionality that, in sets, order is irrelevant and there are no multiple occurrences.

For example, $\{1, 2, 3\} = \{3, 2, 1, 2, 1, 3\}$ simply because the elements of both sets are the same.

Extension versus Intension

A more subtle point is the following: the description of the set is also irrelevant, all that matters are the actual elements.

Here is an example of two sets of natural numbers:

$$A = \{1, 2\}$$

$$B = \{n \in \mathbb{N}^+ \mid x^n + y^n = z^n \text{ has solution in } \mathbb{N}^+ \}$$

Then $A = B$, but this is Fermat's Last "Theorem" and requires a very complicated, non-elementary proof (at least at present no one knows of a short, elementary proof and there are good reasons to believe that none exists).

So equality of sets can be exceedingly complicated even when one of the sets in question is finite.

Subsets

Definition 3. x is a **subset** of y if

$$x \subseteq y \text{ if } \forall z (z \in x \rightarrow z \in y).$$

x is a **proper subset** of y , $x \subset y$ if $x \subseteq y \wedge x \neq y$

Thus $x = y \iff x \subseteq y \wedge y \subseteq x$.

Example 4. For any set S we always have $\emptyset \subseteq S$.

Example 5. For arithmetic types we have $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

The last example may seem blindingly obvious and boring, but it is actually quite tricky. There are many reasonable ways to define these arithmetic sets of numbers so that the inclusions do not hold.

Integers

To see where this problem comes from, suppose we have the natural numbers as urelements and we also have addition on them. We can define the integers as follows: Let ρ be the equivalence relation on $\mathbb{N} \times \mathbb{N}$ given by

$$(a, b) \rho (c, d) \iff a + d = c + b.$$

Then the integers can be defined as the quotient structure

$$\mathbb{Z} = \mathbb{N} \times \mathbb{N} / \rho.$$

Of course, according to this definition \mathbb{Z} is not a subset of \mathbb{N} .

But we can identify \mathbb{N} with such a subset: the $n \geq 0$ is identified with

$$\{ (a, b) \in \mathbb{N} \times \mathbb{N} \mid a = b + n \}.$$

Painful, but very precise. And quite similar to what happens in programming.

Induction and Subsets

We can easily express the Induction Principle on \mathbb{N} using subsets:

Lemma 2. *Let $A \subseteq \mathbb{N}$. Suppose $0 \in A$ and for all x : $x \in A \rightarrow (x + 1) \in A$. Then $A = \mathbb{N}$.*

Proof.

Suppose otherwise, so $A \neq \mathbb{N}$. So there exists an x , $x \notin A$. Pick the least $a \notin A$. Clearly $a \neq 0$. But then $a - 1 \in A$, contradicting the second hypothesis.

□

Of course, this is but the tip of an iceberg: many other sets are also generated by applying certain operations (here: the successor function) to given primitive elements (here: 0). They all obey a similar induction principle.

Powerset

Definition 4. The **powerset** of a set is the set of all its subsets.

Notation:

$$\mathfrak{P}(y) = \{ x \mid x \subseteq y \}$$

Example 6.

$$\mathfrak{P}(\emptyset) = \{ \emptyset \}$$

$$\mathfrak{P}(\{a\}) = \{ \emptyset, \{a\} \}$$

$$\mathfrak{P}(\{a, b\}) = \{ \emptyset, \{a\}, \{b\}, \{a, b\} \}$$

Lemma 3. If A has n elements, then $\mathfrak{P}(A)$ has 2^n elements.

Proof by Induction

Base case: $n = 0$ is correct.

Induction step: Suppose $A = \{a_1, a_2, \dots, a_{n-1}, a_n\}$ and let $x \subseteq A$.

Case 1: $a_n \notin x$.

So x is a subset of $\{a_1, a_2, \dots, a_{n-1}\}$. There are 2^{n-1} such subsets by IH.

Case 2: $a_n \in x$.

Then $x = \{a_n\} \cup y$ where y is a subset of $\{a_1, a_2, \dots, a_{n-1}\}$. Again, there are 2^{n-1} such subsets by IH.

Total: $2^{n-1} + 2^{n-1} = 2^n$. □

► More interesting (and much harder) is the question: What happens when x is infinite?

Exercises

These are all straightforward applications of the definitions.

Exercise 3. *Show that symmetric difference is associative: $A \Delta (B \Delta C) = (A \Delta B) \Delta C$.*

Exercise 4. *Show that $A \subseteq B$ implies that $A \Delta (B - A) = B$.*

Exercise 5. *Show that $A, B \subseteq C$ implies that $(C - A) \Delta B = C - (A \Delta B)$.*

Exercise 6. *Show that $A \subseteq B$ iff $B - (B - A) = A$.*

Exercise 7. *Show that $\text{pow}(A \cap B) = \text{pow}(A) \cap \text{pow}(B)$. How about union instead of intersection?*

Frege's Ghost

Frege's Ghost

You wake up one morning to find that Frege's Ghost (see below for the real Frege) has ruined your favorite C++ compiler: it will not support any data type other than set.

You urgently need to implement lists of integers for a research project. But sets are unordered and there are no integers lying around.

Can we somehow still implement integer lists on our broken compiler?

► Surprisingly, the answer is a resounding **Yes**.

Of course, we have to assume that control structures still work, and that the compiler provides some basic operations on sets. Given such minimal support, we can implement integers and lists as pure sets.

This is just the tip of an iceberg, in mathematics and computer science “everything” can be implemented as a set.

Pairs

First let us deal with the issue of representing lists as sets.

Definition 5. *The (Kuratowski) pair of x and y is*

$$(x, y) = \{\{x\}, \{x, y\}\}$$

Lemma 4. *$(u, v) = (x, y)$ implies $u = x$ and $v = y$.*

Careful, we are dealing with sets: $(x, x) = \{\{x\}\}$.

Exercise 8. *Prove the lemma.*

Exercise 9. *Find another way to implement pairs as sets.*

Exercise 10. *Does the attempt $(x, y) = \{x, \{x, y\}\}$ succeed?*

Tuples aka Lists

Pairs can be extended to lists, here usually called *n-tuples*, by induction. For $n \geq 3$ set

$$(a_1, \dots, a_n) = (a_1, (a_2, \dots, a_n)).$$

Lemma 5. $(a_1, \dots, a_n) = (b_1, \dots, b_n)$ implies $a_i = b_i$.

Note, though, that we cannot weaken the hypothesis to $(a_1, \dots, a_n) = (b_1, \dots, b_m)$.

Exercise 11. *Explain what goes wrong and how to fix it. (Assume we already know how to define natural numbers.)*

Exercise 12. *Find an alternative way to implement lists as functions.*

Frege's Ghost Creates Numbers

Can we implement natural numbers? No problem. Use the successor function (union is built-in)

$$S(x) = x \cup \{x\}$$

Represent a natural number n by a set \underline{n} as follows:

$$\underline{0} \rightsquigarrow \emptyset$$

$$\underline{n} \rightsquigarrow \underbrace{S(S(\dots S(\emptyset) \dots))}_n$$

So, $\underline{3} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ represents the number 3.

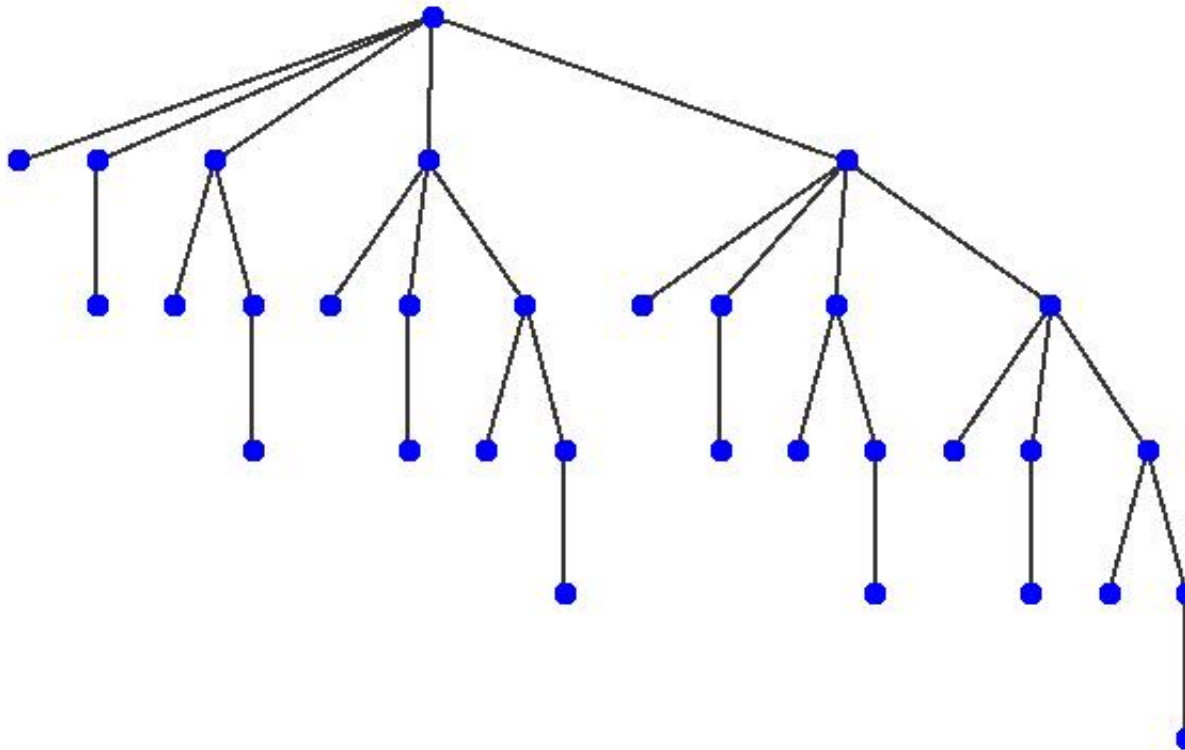
Definition 6. *These sets are the (finite) von Neumann ordinals N_n .*

There are also infinite von Neumann ordinals, more on this later.

von Neumann Trees

Note that we can represent any pure set by a tree: the leaves are instances of the empty set, and lines indicate membership. The internal nodes are the compound sets used to construct the main set which is represented by the root.

Here is the picture for N_5 .



Digression

The previous pictures are indicative of the serious flaw of the set-theoretic approach to mathematical life that we already mentioned: while it is possible to express (essentially) all objects of mathematical discourse as pure sets, the sets in question are often mind-numbingly complicated.

The “pure set” approach has to be tempered with good intuition, and/or devices such as urelements, otherwise one quickly gets lost in the details.

In particular in the case of arithmetic we could simply assume that $\mathbb{N} = \{0, 1, \dots, n, \dots\}$ is given as a set of urelements.

But if we like, we can replace the urelement n by the von Neumann ordinal N_n and deal with pure sets instead.

This is no more than a change in the level of abstraction; the bread and butter of any computer scientist.

Operations

A data type by itself is useless, we need operations.

For von Neumann ordinals some basic operations are the following.

- $N_x = \emptyset$ iff $x = \emptyset$
- $N_x \in N_y$ iff $x < y$
- $N_x \cup N_y = N_{\max(x,y)}$
- Addition:

$$\text{add}(x, N_0) = x$$

$$\text{add}(x, S(N_y)) = S(\text{add}(x, N_y))$$

Likewise could get multiplication, exponentiation, prime decomposition, any computable function on the integers.

Note that the Frege compiler supports recursion.

And, since efficiency is not an issue, this is a great implementation of integer arithmetic. Beats any Cray, no problem.

An Alternative

Why not simply represent natural numbers as deeply nested empty sets:

$$\underline{0} \rightsquigarrow \emptyset$$

$$\underline{n} \rightsquigarrow \underbrace{\{\dots\{\emptyset}\dots\}}_n$$

In other words, $M_0 = \emptyset$ and $M_{n+1} = \{M_n\}$.

There is nothing fundamentally wrong with this approach (which is due to Zermelo), we could still implement order – but not as elegantly as with von Neumann ordinals.

The real problem is that this approach does not generalize gracefully to infinite numbers (which is important to e.g. to prove termination of nested recursions). The von Neumann ordinals can be generalized very nicely and naturally to transfinite numbers.

Another Alternative

Frege proposed a more ingenious way to capture the naturals. The brilliant idea is to associate n with the collection of all sets of size n .

$$\underline{0} \rightsquigarrow \{\emptyset\}$$

$$\underline{n} \rightsquigarrow \{x \mid \exists y \in \underline{n-1}, a \notin y (x = y \cup \{a\})\}$$

Alas, this approach produces proper classes, not sets, and can't be used directly in this form: a class is a collection of objects that is so large that we cannot quite treat it the same way as ordinary sets (e.g., there is no way we can assign a cardinality to a class).

This problem can be fixed by selecting only a few sets of cardinality n to represent n (the first to appear in some natural hierarchy of sets).

Classes and Sets

The existence of proper classes is a technical difficulty that arises in many formalizations of set theory: some perfectly reasonable collections such as the collection of all sets, all vector spaces, all cardinals, all total orders, and so, are too large to be classified as “sets”.

There are ways to deal with these problems systematically; in particular Gödel-Bernays-von Neumann set theory is an often used system. In a nutshell, one allows for classes whose elements are sets, but classes cannot be nested themselves. Formation is unrestricted

$$X = \{ x \mid \varphi(x) \}$$

but x is required to range over sets whereas X may be a class (and φ cannot quantify over class variables).

For example, Russell’s paradoxical $\{ x \mid x \notin x \}$ is a class, so no contradiction ensues.

Products and Sums

Cartesian Product

Here is a more serious application of Kuratowski pairs.

Definition 7. The **Cartesian Product** of two sets A and B is

$$A \times B = \{ (a, b) \mid a \in A, b \in B \}$$

Again we can extend the pairing operation by induction to n sets:

$$A_1 \times A_2 \times \dots \times A_n = (A_1 \times A_2 \times \dots \times A_{n-1}) \times A_n$$

Example 7.

$$[2] \times [3] = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\}$$

In geometry, the plane can be thought of as $\mathbb{R} \times \mathbb{R}$.

And ordinary 3-space is $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

A Nuisance

We have not really said what a Cartesian product of length 0 or 1 should be; since we are dealing with some kind of product this is inelegant at best.

In analogy to multiplication we would expect something like

$$\prod_{i=1}^0 A_i = \mathbf{1}$$
$$\prod_{i=1}^1 A_i = A_1$$

The question is what should $\mathbf{1}$ be? We are dealing with sets, not numbers.

Exercise 13. *Come up with a common sense solution for these problems.*

Families

One often needs to refer to a collection of sets that are “numbered” by some index set. It turns out that it is best to keep things general and allow arbitrary index sets.

Definition 8. Let I be a set, the **index set**. A **family** of sets, indexed by I is a map A with domain I .

Notation: $(A_i)_{i \in I}$ to indicate that $A(i) = A_i$.

This should be perfectly familiar when I is one of the usual choices for index sets: $[n] = \{1, 2, \dots, n\}$, $(n) = \{0, 1, \dots, n - 1\}$, \mathbb{N} , \mathbb{Z} or \mathbb{R} .

Families usually appear in conjunction with infinitary operations. For example, we may consider the family of closed real intervals $A_n = [0, 1/n] \subseteq \mathbb{R}$ for $n \geq 1$. Then $\bigcap_n A_n = \{0\}$. On the other hand, for the open intervals $A_n = (0, 1/n) \subseteq \mathbb{R}$ we have $\bigcap_n A_n = \emptyset$.

General Products

We know how to form the Cartesian product of two sets. Can we generalize this to a product of a whole family of sets?

Definition 9. Let $(A_i)_{i \in I}$ be a family of sets and set $A = \bigcup_i A_i$. The **Cartesian product** of this family is defined by

$$\prod_{i \in I} A_i = \{ \alpha : I \rightarrow A \mid \alpha(i) \in A_i \}.$$

Thus, $\prod_{i \in I} A_i$ consists of all I -indexed sequences of elements in $\bigcup_i A_i$ subject to the condition that the i th element has to be taken from A_i .

For example, if $I = \mathbb{N}$ and $A_i = \mathbf{2}$ then $\prod_{i \in \mathbb{N}} A_i$ is just the collection of all infinite binary sequences. This would often be written simply as $\prod_{\mathbb{N}} \mathbf{2}$.

Likewise, all real sequences (a central notion in analysis) can be obtained as $\prod_{\mathbb{N}} \mathbb{R}$.

Universal Property

So what's so special about products? It turns out the crucial property of products $A = \prod A_i$ can be summarized rather succinctly.

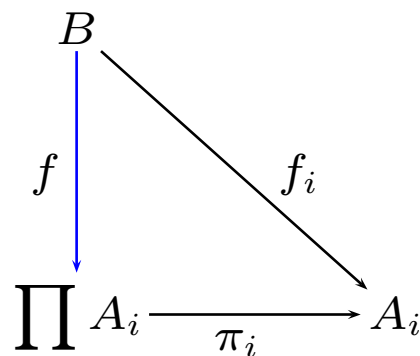
Let

$$\pi_j : \prod A_i \rightarrow A_j \quad \pi_j(\alpha) = \alpha(j)$$

be the projection map onto the j th component.

Theorem 1. *Universal Property*

Given a set B and a family of functions $f_i : B \rightarrow A_i$ there is a unique function $f : B \rightarrow \prod A_i$ such that $\pi_i \circ f = f_i$ for all $i \in I$.



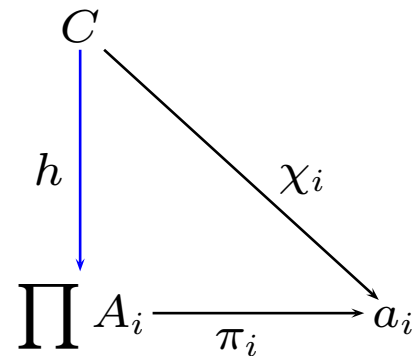
Uniqueness

This theorem is easy to prove: just set $f(b)(i) = f_i(b)$.

What's more important is that any other set C that pretends to be a product must already be the same as $\prod A_i$. By this we mean the following.

Theorem 2. *Uniqueness*

Suppose a set C together with a family of maps $\chi_j : C \rightarrow A_j$ satisfies the same universality property as $\prod A_i$. Then there exists a bijection $h : C \rightarrow \prod A_i$ such that



Whaddayasay?

The Uniqueness theorem may seem a bit abstract, but what it means concretely is that it does not matter exactly how we code up the product, any other reasonable coding will be essentially the same.

Reasonable here means that the alternative product would have to satisfy the same universal property. The bijection simply associates every element in our product with a corresponding element in the alternative product.

This should sound eminently familiar to the computer science student: an array of length n , a list of length n and a record with n fields are all the “same” in some sense, though, of course, there are important differences, too.

Exercise 14. *Show how to reconcile our two definitions of Cartesian product for index sets of size 2. How does this pertain to Cartesian products of arbitrary finite length?*

Disjoint Unions

There is an analogous result for disjoint unions

$$\sum A_i = \{ (a, i) \mid a \in A_i \}$$

together with the injections $\iota_j : A_j \rightarrow \sum A_i$.

Exercise 15. *Explain what the universal property of $\sum A_i$ is.*

Exercise 16. *Show that $\sum A_i$ is unique up to bijections.*

Cantor

Cardinality

Definition 10. *The size of a set is called its **cardinality**.*

Needless to say, this is not much of a definition. We'll have more to say about this later, for the time being use your intuition.

At least for finite sets it is easy to make sense out of this idea: just count the elements. So

$$S = \{a_1, a_2, \dots, a_{n-1}, a_n\}$$

has cardinality n .

Of course, we have tacitly assumed here that the representation of S in the curly braces does not have repetitions; all the a_i must be distinct.

Cardinality

We write $|S| = n$ or sometimes $\text{card}(S) = n$ for the cardinal number of S .

Example 8.

$$|S| = 0 \iff S = \emptyset$$

$$|[n]| = n$$

$$|[n] \times [m]| = n \cdot m$$

$$|[n] \times [n] \times [n]| = n^3$$

It can be amazingly difficult to determine the sizes of various finite sets; the field of combinatorics has developed a rich collection of tools for this purpose.

More Cardinality

Suppose A and B are finite. What is

$$|A \cup B| = ???$$

Not just $|A| + |B|$: that only works when $A \cap B = \emptyset$.

In general we must correct the over-count:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Exercise 17. Find the right expression for $|A \cup B \cup C|$.

Exercise 18. Generalize to n sets A_1, \dots, A_n .

Formalization

So, Where's the Problem?

So far so good. It looks like we can handle sets just fine informally, using common sense.

In mathematics, sets were first studied systematically by Georg Cantor (1845-1918). Surprisingly, he was lead to the study of set during his wok on Fourier analysis.

Cantor found lots of strange properties of infinite sets, and at the time many mathematicians disagreed about the nature of infinite sets.

This led Gottlob Frege (1848-1925) to propose an *axiom system* for sets, which was supposed to settle all debate once and for all.

His system is a bit different from its modern counterparts (see the lecture on Formal Systems). To simplify matters we will reinterpret it slightly and express the axioms in modern terminology.

Frege's Axioms

Here is a modern re-write of G. Frege's axiom V which states that two sets are equal if, and only if, they contain the same elements.

Since Frege also assumes that every concept P is associated with an "extension", essentially the set $\{z \mid P(z)\}$ of all objects with property P , we also have a powerful set formation axiom as below.

- **Extensionality**

$$x = y \text{ if } \forall z (z \in x \iff z \in y)$$

- **Formation**

For any property $P(z)$:

$$\exists x \forall z (z \in x \iff P(z))$$

For us, the quantifiers all range over the collection of all sets. Note that by Extensionality the set x in Formation is unique.

Enormous Power

It would be a slightly tedious to carry out all the details, but one could give a completely set-theoretic treatment of, say, calculus, using just these two axioms.

Starting from the von Neumann ordinals as representations of the natural numbers one proceeds in stages and constructs the integers, the rationals and the reals.

Functions on the reals are then sets of pairs of reals, and so on and so on.

Frege's two axioms (and a lot of stamina) are the only tools needed to do this.

Exercise 19. *Find a way to express integers in terms of sets of von Neumann ordinals.*

Russell's Paradox

Unfortunately, there is a fatal flaw in Frege's system: his axioms are inconsistent.

Bertrand Russell pointed this out in a letter to Frege in 1902 (just when Frege was getting ready to publish the second volume of his Grundgesetze).

Russell's Set

$$S = \{ z \mid z \notin z \}$$

Looks strange, but why not?

But then both $S \in S$ and $S \notin S$ lead to a contradiction, and we're sunk.

$$\frac{S \in S \vee S \notin S}{\perp} \quad \frac{\frac{[S \in S]}{S \notin S} df \quad \frac{[S \notin S]}{S \in S} df}{\perp} (\neg e) \quad \frac{\perp}{\perp} (\vee e)$$

Who Cares?

How serious is this, really? Well, it nearly killed Frege . . . He mentioned Russell's result in an appendix, and presented a supposed remedy (which Russell originally accepted, but later both he and Frege realized that the remedy did not work).

However, unless you decide to become a logician (and thus permanently unemployable), you will never encounter inconsistent set constructions.

There are ways to fix this problem, but they are a bit technical, and not needed for standard applications. As long as you apply common sense to set formation, you'll be OK.

Here is a glimpse at a system of axioms that works, and that has become the de facto standard: Zermelo-Fränkel set theory.

It is a truly remarkable feature of this system that it is powerful enough to express most of mathematics in a clean and precise way. Yet, it uses only a handful of axioms (well, really schemata) that are fairly easy to accept intuitively.

Zermelo-Fränkel Set Theory

One way to avoid Russell's paradox is to write down a list of carefully designed axioms that describe all properties of sets and are very conservative when it comes to set formation.

They keep Frege's Extensionality axiom.

$$x = y \iff \forall z (z \in x \iff z \in y)$$

And add a number of modest *set-existence axioms*: axioms that say sets with certain narrow properties exist – not like Frege's sledgehammer Formation axiom.

E.g., there is an axiom that says “the empty set exists”.

Then there is one for unordered pairs $\{x, y\}$, for union, and so on.

Needless to say, this is much more tedious than Frege's Formation, but it gets us around Russell.

We'll just look at a few, the others are on the web.

Basic Axioms

Empty Set

$$\exists u \forall z (z \notin u)$$

Unordered Pair

$$\exists u \forall z (z \in u \Leftrightarrow z = x \vee z = y)$$

Union

$$\exists u \forall z (z \in u \Leftrightarrow \exists y (y \in x \wedge z \in y))$$

Comprehension (Aussonderungsaxiom)

$$\exists u \forall z (z \in u \Leftrightarrow z \in x \wedge P(z))$$

So Unordered Pair says: $\{x, y\}$ exists.

More Axioms

The *successor function* is defined by $S(x) = x \cup \{x\}$.

Power Set

$$\exists u \forall z (z \in u \Leftrightarrow z \subseteq x)$$

Infinity

$$\exists u (\emptyset \in u \wedge \forall z (z \in u \Rightarrow S(z) \in u))$$

Foundation

$$\exists u (u \cap x = \emptyset)$$

A Weird Axiom

The last axiom deals with applying functions to a set of elements. A property Q is called a *functional property* if

$$\forall w, u, v (Q(w, u) \wedge Q(w, v) \rightarrow u = v)$$

For any functional property Q we have the following axiom:

Replacement Axiom

$$\exists u \forall z (z \in u \leftrightarrow \exists w (w \in x \wedge Q(w, z)))$$

The replacement axiom says in effect that we can apply a function to a set, and get back another set. This axiom does not follow from the previous ones, it adds additional strength to the system.

Axiom Schmaxiom

Note that the axiomatic approach to set theory makes no attempt whatsoever to define what a set is or to define the membership relation. Rather, we pin down a number of basic properties of the objects in the universe and the binary relation on them.

So, axioms may appear trivial if one thinks about the interpretation as sets, but that's in the eye of the beholder.

Also note that some axioms are not quite so obvious. The most important one is the *Axiom of Choice (AC)*, which was introduced by Ernst Zermelo in 1904 to give a construction of a well-order of the reals.

Zermelo's idea was not too well-received at the time, but it has become mainstream by now: the Axiom of Choice is indispensable for many arguments in mathematics.

Alas, as we will see, it also has some nearly absurd consequences.

Axiom of Choice

Suppose we have a collection $(A_i)_{i \in I}$ of non-empty sets that are pairwise disjoint:
 $i \neq j \rightarrow A_i \cap A_j = \emptyset$.

We would like to construct a *choice set* C for A_i : a set that selects exactly one element from each A_i .

$$C \cap A_i = \{a_i\} \quad \text{for all } i.$$

We can also think of this as a *choice function*, a map

$$C : I \rightarrow \bigcup A_i \quad C(i) \in A_i$$

(in which case we can omit the disjointness condition).

If I is finite, or if the sets A_i carry a nice structure such as being subsets of \mathbb{N} then (AC) is not needed: the other axioms guarantee the existence of choice sets/functions.

But in general the existence of a choice set/function does not follow from the other axioms.

(ZFC)

In many ways (AC) is so natural that one feels it ought simply to be incorporated in the standard system of set theory (some logicians would furiously disagree).

Definition 11. *Zermelo-Fränkel set theory (ZF) is defined by the axioms above. (ZFC) is (ZF) plus the Axiom of Choice.*

It is good practice to point out whenever a result depends on (AC) rather than just use it tacitly. For example, the famous Nielsen-Schreier theorem that states that a subgroup of a free group is again free requires (AC), as does the “fact” that the additive group of the reals is isomorphic to the additive group of the complex numbers.

AC and Well-Orders

Here is another important application, a positive answer to the question whether we can perform induction on the real numbers?

In order to do induction we need to well-order \mathbb{R} . Clearly, the standard order does not work; it fails already on the integers or the positive rationals.

But how do we define a well-order? It was Zermelo's idea to use the Axiom of Choice to define an abstract ordering on the reals that is provably a well-order. To this end, pick a choice function

$$f : \mathfrak{P}(\mathbb{R}) \rightarrow \mathbb{R}$$

such that $\emptyset \neq X \subseteq \mathbb{R} \rightarrow f(X) \in X$. Then enumerate the reals according to

$$r_0 = f(\mathbb{R})$$

$$r_1 = f(\mathbb{R} - \{r_0\})$$

$$r_2 = f(\mathbb{R} - \{r_0, r_1\})$$

...

A Well-Order

Repeating this process we obtain a well-order $r_0 < r_1 < \dots < r_n < \dots$

But the “...” here is really tricky: since there are more reals than natural numbers we need transfinite induction. After we have exhausted all stages $n \in \mathbb{N}$ we continue:

$$\begin{aligned} r_\omega &= f(\mathbb{R} - \{r_i \mid i \in \mathbb{N}\}) \\ r_{\omega+1} &= f(\mathbb{R} - \{r_i \mid i \in \mathbb{N}\} - \{r_\omega\}) \\ r_{\omega+2} &= f(\mathbb{R} - \{r_i \mid i \in \mathbb{N}\} - \{r_\omega, r_{\omega+1}\}) \\ &\dots \end{aligned}$$

Note the ω hiding in the subscript. This is an *ordinal number*, an extension of the natural numbers for purposes of enumeration. There is a nice definition of these objects in terms of von Neumann ordinals. See the notes on ordinals and cardinals for details.

Why Bother?

If (AC) is useful, and makes intuitive sense, why not just adopt it and not make a big fuss about it?

Because (AC) also has a dark side, a few strange consequences, and a few extremely bizarre consequences.

First off, (AC) implies that there are sets of reals that fail to be Lebesgue measurable. This is certainly a bit counter-intuitive; it is not clear what exactly should prevent us from assigning a measure to an arbitrary set of reals.

Worse, a result by F. Hausdorff from 1914 states that one can partition a sphere (after removing countably many points) into three parts A , B and C such that all three pieces are congruent, and are also congruent to $B \cup C$.

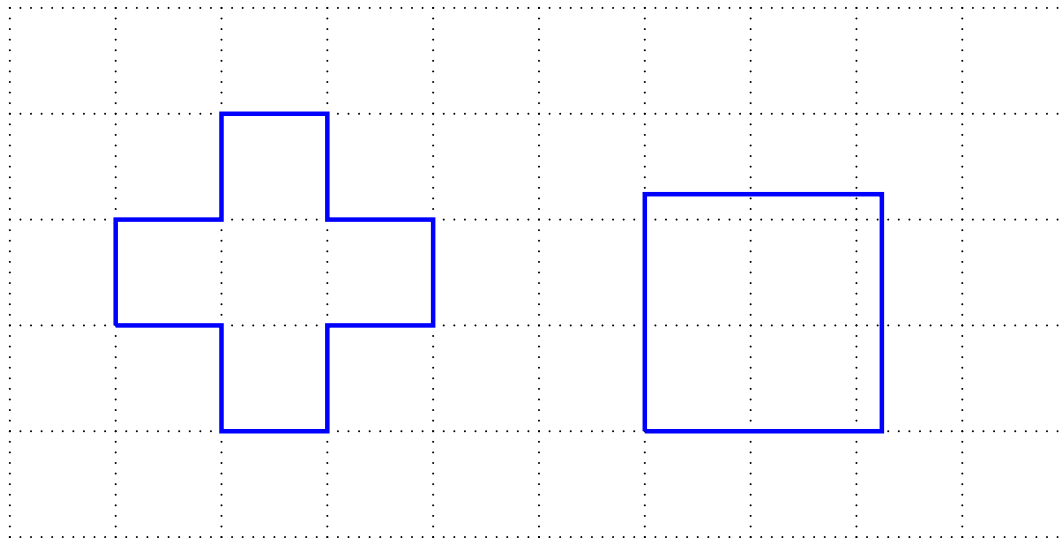
This sounds more like a paradox than a theorem.

Cutting Things Up

By comparison, consider the following entirely reasonable and unsurprising result in the plane.

Theorem 3. *Bolyai-Gerwin Theorem*

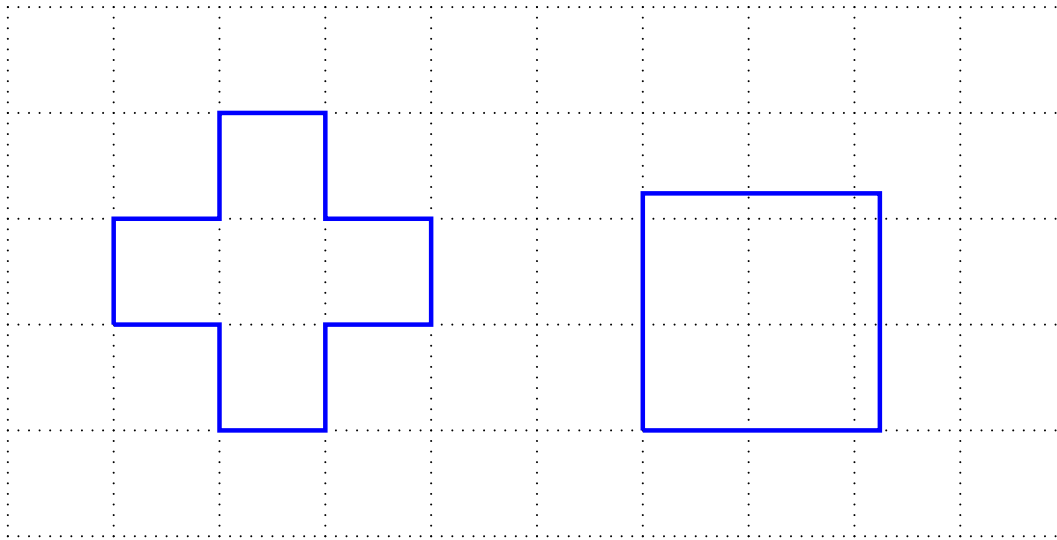
Let P and Q be two polygons of equal area. Then P can be partitioned into finitely many triangles that can be reassembled to form Q .



The square on the right has size $\sqrt{5}$ by $\sqrt{5}$ and thus area 5, just like the cross on the left.

Challenge

Take a pair of (mental) scissors, and find a simple decomposition of the cross on the left whose pieces can be rearranged to form of the square on the right.



The fewer cuts, the better.

The Banach-Tarski Paradox

Of course, the restriction to triangles is severe. Also, having another dimension helps. Still, the next theorem which builds on Hausdorff's result is truly wild and sounds positively like a serious error. Needless to say, its proof requires the Axiom of Choice.

Theorem 4. *Banach-Tarski Paradox*

The unit sphere can be decomposed into finitely many pieces, that can be reassembled to form a sphere of radius 2.

The pieces in the Banach-Tarski decomposition are very strange and cannot be visualized. In particular we cannot assign qualities such as "volume" to these pieces: otherwise we would immediately have a contradiction.

So, this result is a correct theorem of (ZFC), but it is very, very counter-intuitive. In a sense, this is much worse than Cantor's problems with cardinality.

The Measure Problem

We would like to find a d -dimensional measure, a map $\mu : \mathfrak{P}(\mathbb{R}^d) \rightarrow \mathbb{R}_0^+$ such that

- Any two equidecomposable sets $A, B \subseteq \mathbb{R}^d$ have the same measure: $\mu(A) = \mu(B)$.
- The measure is additive on disjoint sets:
 $A \cap B = \emptyset$ implies $\mu(A \cup B) = \mu(A) + \mu(B)$.
- The measure is normalized: $\mu([0, 1]^d) = 1$.

Theorem 5. *Hausdorff 1914*

Measures do not exist for dimensions $d \geq 3$.

Theorem 6. *Banach 1923*

Assuming the Axiom of Choice, measures do exist for dimensions $d = 1, 2$.

Independence

Theorem 7. *Gödel 1938*
ZFC is equiconsistent with ZF.

Theorem 8. *Cohen 1963*
(AC) is independent of ZF.

Theorem 9. *Solovay 1970*
There is a model of ZF where every set of reals is measurable.

So the Banach-Tarski paradox depends crucially on the Axiom of Choice.

Summary

- Intuitive set theory leads to paradoxes, but the problematic constructions are far removed from “real world” applications.
- Axiomatic set theory can deal with paradoxes, but at the cost of added technical difficulties.
- In practice, intuitive set theory, augmented with a bit of axiomatics, serves as a solid foundation of mathematics and computer science.
- In particular elementary concepts such as relations, functions, computable functions and so on can all be defined within set theory.