Set theory for intro discrete math, Fall 2024 version

Randall Holmes

Sept 11 2024: using this, since my impromptu lecture on Monday matched pp. 1-8 quite well

1 Introduction

This document discusses simple set theory foundations for discrete math. It is meant to be read. Not everything in it is test material; if something scares you, ask whether it might be on a test.

I do believe that some axioms and definitions are appropriate, and I give some. But I try to restrict them to what is actually useful for discrete math at a beginning undergraduate level.

Something interesting happens (which is part of the reason I am writing this). The sum of what we teach in a discrete math course does add up to an axiomatic set theory, close to Zermelo set theory though not identical to it, and adequate for the foundations of classical mathematics.

The idea here is to exhibit situations in discrete math where the set and function abstractions come up naturally, and develop the exact axioms and definitions that are appropriate in this context.

Natural finite examples of the set and function abstractions come up in combinatorics.

When I ask, how many subcommittees of three people can be formed from a committee of ten people, I am actually asking, how many subsets of size three does a subset of size ten have?

And there is something disingenuous here...of course, two subcommittees can have the same roster. We are asking a question about sets in the background so we want to presume the identity condition for sets...two sets are the same iff they have the same elements. A semi-real-world question along

the same lines which avoids this issue is...a subcommittee of three is to be formed from a committee of ten people...how many possible rosters of members are there? Of course we know that the order of the members doesn't matter – but this is another way one could become confused as to what the correct answer is here if the set concept isn't familiar.

Another thing to notice in any discrete math application is that the size (at least of a finite set) is an important part of the interface of the data type of sets (at least, finite sets). A lot can be done with just the notions of membership and equality, but in discrete math we are interested in counting.

2 Sets

2.1 Primitive notions and equality of sets

I shall begin with the simplest assumptions about sets.

primitive properties and relations: Some objects in our world are *sets*.

Primitive notion 1 (sets and individuals): Some objects are *sets*. Objects which are not sets we call *individuals*[1].

Primitive notion 2 (membership): There is a relation of *membership*, written $\in$, which holds between general objects and sets they "belong to".

Primitive notion 3 (equality): We presume familiarity with the general notion of equality, and with its basic logical properties ($x = x$ [equality is reflexive], and "if $x = y$ and $P[x]$ then $P[y]$" [substitution of equals for equals]).)

We say "x is an element of a" or "x is a member of a", "x belongs to A" or "x is contained in a" to mean $x \in a$. "x is included in a" has another meaning for us. We write $x \notin A$ for "x is not an element of A".

We try to avoid saying "x is in A", because this is ambiguous: it can be confused with the subset relation.

identity criterion: With any data type, we want to be able to tell when two objects of that type are the same.

If A and B are sets, $A = B$ holds if and only if A and B have the same elements, that is, for any x, $x \in A$ holds exactly when $x \in B$ holds. Another way of putting this is, There is no element of A which does not belong to B, and there is no element of B which does not belong to A. This avoids vacuous quantification.

This is summarized in

[1] what I call "individuals" have also been called "atoms" or "urelements"

Axiom 1 (extensionality): If A and B are sets, and no element of B is not an element of A, and no element of A is not an element of B, then $A = B$.

Equivalently, $((\forall x \in A : x \in B) \wedge (\forall x \in B : x \in A) \rightarrow A = B$.

individuals and empty set: From the introduction of the membership relation, we extract this statement:

Axiom 1b (axiom of sethood): If $x \in A$, then A is a set. Equivalently, if A is an individual and x is any object, $x \notin A$ (in which form we would call it the axiom of individuals).

We call this axiom 1b because it is a footnote to extensionality, and also because usual treatments of set theory do not allow for individuals at all (but this tends not to be convenient in undergraduate discrete math classes).

A special case of the identity criterion is that if A and B are sets, and both have no elements at all, then they are equal. There is at most one set with no elements (we will introduce an axiom that says there is one in a moment). In addition, any objects in our world that are not sets have no elements. These objects can be called by various names: we call them individuals.

An important point in mathematics pedagogy is that official foundations of mathematics usually say that everything is a set, but common sense allows for individuals. Moreover, familiar mathematical objects (such as the natural numbers) have implementations as sets, but there is nothing inevitable about these implementations, and it is often natural in an undergraduate discrete math text to treat items not explicitly given as sets as individuals (natural and real numbers, for example).

2.2 Axioms which allow us to construct sets; notation for sets

properties define sets: If we have a set A given and a statement $P[x]$ about objects x in general, there is a set $\{x \in A : P[x]\}$ for which this is true: for any object a, $a \in \{x \in A : P[x]\}$ if and only if $a \in A$ and $P[a]$.

The notation "$\{x \in A : P[x]\}$" is called set builder notation: it is important, and it has variations which are important.

This says: Given a set A and a property P, we can extract the collection of all elements of A which have the property P as a new set.

Notice that for any A, $\{x \in A : x \neq x\}$ is an empty set, and there is only one, for which we adopt the notation $\emptyset$.

We codify this as an axiom.

Axiom 2 (separation): For any sentence $P[x]$ about general objects x, and any set A, we have an object $\{x \in A : P[x]\}$, which is a set, and the axiom "for any a, $a \in \{x \in A : P[x]\}$ if and only if $a \in A$ and $P[a]$".

subset relation and power set: We define $A \subseteq B$, read "A is a subset of B" as "A is a set and B is a set and for any x, if x is an element of A, x is an element of B" **or** "A is a set, B is a set, and anything which is not an element of B is not an element of A". The contrapositive formula has value because the implicit quantifier is never vacuous: there is always something not in B, and it's clear from this definition that $\emptyset \subseteq B$: anything not in B is not in $\emptyset$.

Equivalently, we can define $A \subseteq B$ as $(\forall x \in A : x \in B)$.

Note that $\emptyset \subseteq A$ and $A \subseteq A$ always hold.

We assert as a basic assumption that

Axiom 3 (power set): for any set A, there is a set $\mathcal{P}(A)$, called the power set of A, whose elements are exactly the subsets of A.

Notice that the axiom of extensionality actually says that if $A \subseteq B$ and $B \subseteq A$, then $A = B$. That is exactly what it says.

list notation for finite sets: The notation $\{x\}$ denotes the set whose only element is x.

The notation $\{x, y\}$ denotes the set whose only elements are x and y.

The notation $\{x, y, z\}$ denotes the set whose only elements are x, y, z.

And so forth. List notation $\{x_1, x_2, \ldots, x_n\}$ defines a set whose only elements are the x_i's. The general definition of this (familiar) list notation is technically exacting to state (we might do it eventually).

Notice that order and repetitions of items do not make any difference in the reference of this notation. $\{x, x\}$ is the same set as $\{x\}$. Notice that this means that you cannot tell that a set written $\{x, y\}$ has two elements unless you are given the information that x and y are distinct.

$\{x, y\}, \{y, x\}, \{x, y, y\}$ etc. are all the same set.

assumptions behind list notation: Behind the ability to write this notation, there are two basic assumptions.

Axiom 4 (singletons): For any object a, there is a set $\{a\}$ such that for any x, $x \in \{a\}$ exactly if $x = a$.

It is fun to notice that if a is a set, $\{a\} = \{x \in \mathcal{P}(a) : x = a\}$ is already given by axioms previously stated. But we are not presuming that every object is a set, so we need the axiom of singletons. [2]

Axiom 5 (binary union): For any sets A and B, there is a set $A \cup B$ such that for any x, $x \in A \cup B$ if and only if either $x \in A$, or $x \in B$, or both. This set is called the union of A and B.

Now $\{x, y\} = \{x\} \cup \{y\}$ and more generally $\{x_1, x_2, \ldots, x_n\} = \{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_n\}$.

[2] An inverse operation of sorts to the singleton operation is definite description. It would be handy to have an operator θ such that $\theta(\{x\}) = x$ and $\theta(u) = \emptyset$ if x is not a singleton set. Then $\theta(\{x \in A : P[x]\}$ would represent the unique x in A such that $P[x]$ if there is one. We note the desirability of this without postulating it. We would read $\theta(A)$ as "the unique element of A", and use of this notation would usually presume that A has exactly one element.

Relations of part and whole on sets?: There is a temptation to say that a set is a whole made up of its elements. A classic textbook written by a man who certainly knew better gave packs of wolves and bunches of grapes as examples of sets.

This temptation should be firmly resisted. If A is part of B and B is part of C, then A is part of C, for any objects A, B, C on any reasonable understanding of the relation of part to whole. But if a, b are any two distinct objects, $a \in \{a, b\}, b \in \{a, b\}$, and $\{a, b\} \in \{\{a, b\}\}$. If membership were transitive as the relation of part to whole is, it would follow that $a \in \{\{a, b\}\}$ and $b \in \{\{a, b\}\}$. But $\{\{a, b\}\}$ has only one element, the set $\{a, b\}$: a and b are not both in it (even if one of them were weirdly the same as $\{a, b\}$)

Related temptations are the common desire to say that $\emptyset$ belongs to every set, or to write $\emptyset$ as $\{\emptyset\}$ (the latter is a set with one element while $\emptyset$ has no elements).

It is important to notice that none of this has to do with famous worries about infinite sets: these issues arise from the simplest construction of sets by finite listing. They do have to do essentially with allowing sets to be elements of sets, which is an important move for the uses of sets intended in mathematics.

The natural relation of part to whole on sets is the subset relation. We say "A is included in B" for $A \subseteq B$, not $A \in B$. Note that if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

It may seem peculiar that every set has the empty set as a part: a part of a set could be defined as a *nonempty* subset of the set, which would preserve the idea that disjoint sets have no common part (though they do have the common subset $\emptyset$)

Notice that $x \in A$ is equivalent to $\{x\} \subseteq A$: the elements of A correspond to but are not identical with atomic parts (smallest possible nonempty subsets) of A.

Notice that the examples given in the famous textbook are objects with disjoint parts of an understood kind: a pack of wolves does have a natural association with a set of wolves, and a bunch of grapes with a set of grapes. But the mass of all human cells is roughly speaking the same as the mass of all human beings, while the set of human cells is a lot larger then the set of human beings.

Other interesting binary operations on sets: We define $A \cap B$ as

$$\{x \in A : x \in B\}.$$

This is called the intersection of A and B.

Exercise: prove that $\{x \in A : x \in B\} = \{x \in B : x \in A\}$

We define $A - B$ (also sometimes written $A \setminus B$) as $\{x \in A : x \notin B\}$. Here, $x \notin B$ simply means "x is not an element of B". This is called the set difference of A and B or the complement of B relative to A.

These, along with union, are the basic operations for the parlor game of Venn diagrams, which we will play (which does have its uses, mostly to illustrate very simple properties of two or three sets).

Notice that the two operations introduced here require no new axioms, because the set defined is included in a set already given. Also notice that this was not true of unions of two sets, which is why we needed an axiom for that.

2.3 Exercises

Here are some exercises, due on Monday the 17th.

1. Let $A = \{1, 2, 3, 4\}$ and $B = \{1, 3, 5, 7\}$. Give the following sets in the same finite list notation. Compare with Levin 0.3.1.
   (a) $A \cup B$
   (b) $A \cap B$
   (c) $A \setminus B$ (which I might write $A - B$)
   (d) $B \setminus A$
2. This is a favorite kind of test question for me. In each part, supply the missing relation, either $\in$, $\subseteq$, both, or neither.
   (a) $\{1\}\_\_\_\{1, 2, 3\}$
   (b) $-3\_\_\_\mathbb{Z}$
   (c) $\{1, 2\}\_\_\_\{1, \{1, 2\}, 2, 3\}$
   (d) $3\_\_\_\{\{3, 4, 5\}\}$
   (e) $\emptyset\_\_\{\emptyset\}$
3. In this question I'm giving a lot of explanation in English which I would not give on a test question (guiding you in reading notation you might encounter).
   (a) Let $A = \{1, 2, 3, 4\}$. Give $\mathcal{P}(A)$ in list notation (it has sixteen elements, I know).
   (b) Now give $\{x \in \mathcal{P}(A) : |x| = 2\}$ in list notation (this is the collection of elements of the previous set which themselves have two elements).
   (c) State $|\{x \in \mathcal{P}(A) : |x| = 2\}|$. This is the number of elements in the previous set.
   (d) Explain why I do not ask you to give $\mathcal{P}(B)$ in list notation where $B = \{x \in \mathbb{N} : 1 \leq x \leq 20\}$. Hint: how many elements does this set have?
4. Give a Venn diagram proof in the style I did in class (provide keys to shadings and clearly outline the result set in each picture) of

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

   what kind of algebraic property is this?

5. Give a Venn diagram proof in the style I did in class of

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$$

6. Do Levin's problem 0.3.23. Give examples of sets B in list notation which give the largest value of $A \cup B$ and which give the smallest value.

3 Pairs, lists, relations, functions

The abstraction of ordered lists: In addition to considering sets, which are not ordered ($\{x, y\}$ is the same set as $\{y, x\}$) we want to consider ordered pairs (x, y) or ordered lists $[x_1, x_2, \ldots, x_n]$, which are the same exactly if they have the same number of items appearing in the same order (repetitions being significant). Our reasons for using different delimiters for lists will be revealed later.

This is an independent, if related idea. It is interesting that it can be implemented entirely in terms of the set concept. But please notice that there is nothing inevitable or unique about this implementation.

Concrete examples of the general use of the ordered list concept are not hard to come by (permutations versus combinations).

Basic properties of the ordered pair: The basic properties of the ordered pair concept are...

for any objects x, y there is a pair (x, y)

$(x, y) = (z, w)$ if and only if $x = z$ and $y = w$.

A definition of (x, y) as a set which works, and the one which is now almost exclusively used, is $(x, y) = \{\{x\}, \{x, y\}\}$. For this to work as an ordered pair definition, we need to be able to construct it for any x and y (our axioms of singletons and union let us do that) and we need to be able to extract x and y from (x, y) (that is, given an ordered pair we need to be able to identify its first and second components). I tell you how to do this, though this proof and the exact definition of the ordered pair are not examinable content: x is the only object which belongs to all elements of $(x, y) = \{\{x\}, \{x, y\}\}$, and y is the only object which belongs to exactly one element of $(x, y) = \{\{x\}, \{x, y\}\}$. A reason that proofs about this ordered pair notion can be tricky (and a good reminder of the perils of set notation) is that one cannot assume that $\{\{x\}, \{x, y\}\}$ has two elements: if $x = y$, $\{\{x\}, \{x, y\}\} = \{\{x\}\}$.

What we require of the general list concept is

for any $x_1, \ldots, x_n$ there is a list $[x_1, \ldots, x_n]$

$[x_1, \ldots, x_n] = [y_1, \ldots, y_n]$ iff $x_i = y_i$ where $1 \leq i \leq n$.

We actually define lists as functions:

$$[x_1, \ldots, x_n] = \{(i, x_i) : 1 \leq i \leq n\} :$$

our official discussion of functions is below.

The Cartesian product: Given sets A, B, we define the Cartesian product of A and B, which we write $A \times B$, as the collection of all ordered pairs (a, b) with $a \in A$ and $b \in B$.

The existence of the Cartesian product is a consequence of the axioms we have already given.

We define $\mathcal{P}^2(A)$ as $\mathcal{P}(\mathcal{P}(A))$, and more generally $\mathcal{P}^{n+1}(A)$ as $\mathcal{P}(\mathcal{P}^n(A))$ for $n \geq 2$.

It is then straightforward to observe that for any $a \in A$, $b \in B$, both $\{a\}$ and $\{a, b\}$ belong to $\mathcal{P}(A \cup B)$ so $\{\{a\}, \{a, b\}\}$ belongs to $\mathcal{P}^2((A \cup B)$, and so we can define $A \times B$ as the set of all x in $\mathcal{P}^2((A \cup B)$ such that for some $a \in A, b \in B$, $x = \{\{a\}, \{a, b\}\}$:

$$A \times B = \{x \in \mathcal{P}^2(A \cup B) : (\exists a \in A : \exists b \in B : x = (a, b)\}$$

Here we have written the set builder notation very carefully so that you can see that axiom 2 provides us with this set. This might more often be written

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

The expansion above can be taken as a general hint of how to deal with complicated expressions left of the colon in set builder notation (where axiom 2 formally allows only a variable letter and the bounding set) and the bounding set on the left of the colon can be deduced from information given on the right of the colon (clearly $(a, b) \in A \times B$ – once we know that Cartesian products exist in general).

The proofs that Cartesian products exist should not be important directly to you; the mere assertion that they exist should be enough, as you really shouldn't work directly with the details of pairs as sets very much (what you do with them should actually be quite independent of their implementation). What should have some interest is that an implementation is possible!

The definition of a relation: Let A, B be sets. A relation from A to B is a triple $R = (A, B, G)$ where G is a subset of $A \times B$, and where we define (x, y, z) (for this specific purpose) as $((x, y), z)$.

We call A the domain of R ($\mathtt{dom}(R)$), B the codomain of R ($\mathtt{cod}(R)$) and G the graph of R ($\mathtt{graph}(R)$). There is another school of thought (which by temperament I prefer) which identifies a relation with its graph, but there are technical problems with this, because in general the domain and codomain cannot be determined from the graph, and it is common to speak of the domain and codomain as features of the relation.

We define $x\,R\,y$ as the assertion $(x, y) \in \mathtt{graph}(R)$.

We define the image or range of R as the set of y in the codomain of R such that there is x such that $x\,R\,y$.

We define the preimage of R as the set of x in the domain of R such that there is y such that $x\,R\,y$.

Many but not all transitive verbs in mathematics can be read as relations. The most general ones, such as $x = y$, $x \in y$, $x \subseteq y$ cannot, because there are no sets large enough to serve as domain or codomain of these "logical relations".

the definition of functions: A function is a relation $F = (A, B, G)$ with the property that for each $x \in A$, there is exactly one $y \in B$ such that $x\,F\,y$.

We write $f : A \to B$ to mean "f is a function with domain A and codomain B" or, less formally, f is a function from A to B.

This is more concrete than Levin's notion that a function is a rule. As we will see, the set of ordered pairs G codes the rule.

We expand the language a little: it is equivalent and perhaps easier to follow to say that for each $x \in A$, there is $y \in B$ such that $x\,F\,y$ (so the preimage of F is the domain) and for any x, y, z, if $x\,F\,y$ and $x\,F\,z$, then $y = z$.

If F is a function and $x \in \mathtt{dom}(F)$, we define $F(x)$ as the unique y such that $x\,F\,y$. [3]

[3] If we had the definite description operator, we could write

$$F(x) = \theta(\{y \in \mathtt{cod}(F) : x\,F\,y\}).$$

It is worth noting that there are "logical functions" which are not implementable as sets. For example, there can be no function F such that $F(x) = \mathcal{P}(x)$ for all x, since the domain of such a function would be the collection of all sets, which can be shown not to exist by the Russell argument.

You learned a quite different definition of the function concept in high school and in college calculus. We illustrate that our formalization is adequate to support that informal definition.

We set out to define the function $y = 2x + 5$ from real numbers to real numbers. We suppose that we have the set $\mathbb{R}$ of real numbers handy.

We then have the graph of f as the set G of all u in $\mathbb{R} \times \mathbb{R}$ such that there is $x \in \mathbb{R}$ such that $u = (x, 2x + 5)$.

In general, if we are given a definition $y = f(x)$ of a function by a rule, as is usually done in calculus, and we understand a domain A and codomain B of f from context, then

$$f = (A, B, \{(x, y) \in A \times B : y = f(x)\}).$$

A notation for this might be $(x \in A \mapsto f(x) \in B)$. The sets can be omitted if they are understood from context (it is very odd to explicitly provide B as I do here, but I am making a point).

So the function above could be written as $(x \in \mathbb{R} \mapsto 2x+5 \in R)$, which could just be written $(x \in \mathbb{R} \mapsto 2x + 5)$ [certainly] or as $(x \mapsto 2x + 5)$ if you are confident that the domain is understood [$(x \in \mathbb{Z} : 2x + 5)$ is not the same function!]

A weird point which I should mention but not make too much of is that $(x \in \mathbb{R} : x^2 \in \mathbb{R})$ and $(x \in \mathbb{R} : x^2 \in \mathbb{R}^+ \cup \{0\})$ are functions with the same values at the same inputs, but they are not the same function because they have different codomains. They have the same graph: on the alternative view identifying functions with their graphs, they would be the same function.

Images and inverse images: If $f : A \to B$ and $C \subseteq A$, we define $f[A]$ as $\{f(x) : x \in C\}$. which can also be written in the form $\{b \in B : (\exists c \in C : f(c) = b)\}$. We provide the second form to make it clear that

this set exists by axiom 2, and also to provide practice in reading more complicated forms of set builder notation.

If $f : A \to B$ and $C \subseteq B$, we define $f^{-1}[A]$ as $\{a \in A : f(a) \in C\}$. Notice that this involves no mention of the inverse function of f or even any assumption that f has an inverse function (which is fortunate, since we don't mention inverse functions until the next section!)

Levin uses parentheses in these notations. This is traditional (you will see this usage in books) but a bad idea: if there is an object C in the domain of f which is also a subset of the domain of f, which is not at all impossible, then $f(C)$ could mean two different things, and similarly if f has an inverse function and C is an element of the codomain which is also a subset of the codomain, $f^{-1}(C)$ would have two possible interpretations.

Levin also allows $f^{-1}(x)$ to be used to mean $f^{-1}(\{x\})$, and I do not. If f^{-1} exists as an inverse function, this is really bad notation.

Some kinds of function which are commonly considered: A function f is an injection, or one-to-one, if for any $x, y \in \mathtt{dom}(f)$, if $f(x) = f(y)$ then $x = y$.

A function f is a surjection, or onto, if for any $y \in \mathtt{cod}(f)$, there is x such that $y = f(x)$. Notice that in our weird example at the end of the last subsection, the first function is not surjective, and the second is.

A function f is a bijection iff it is one-to-one and onto (i.e., an injection and a surjection).

Inverse relations and functions: For any relation $R = (A, B, G)$, there is an inverse relation $R^{-1} = (B, A, \{(y, x) : (x, y) \in G\})$.

A function f is a relation, so f^{-1} defined as above certainly exists. We say that f has an inverse function exactly if f^{-1} is itself a function. This is true exactly if f is both injective and surjective, that is, iff f is a bijection. So the condition on functions of being a bijection is exactly the same as the condition of having an inverse function.

Our official definition of ordered lists: We decouple lists from ordered pairs. We use the notation $[x_1, x_2, \ldots, x_n]$ to make this clear.

We define an ordered list as a graph of a function whose domain is an interval in the integers (assuming familiarity with the integers; our

treatment below will at least suggest how the integers themselves could fit into our scheme). This allows variations in how they are indexed. If x is an ordered list, x_i is then simply defined as $x(i)$. The list $[x, y, z]$ is for us the set $\{(x, 1), (y, 2), (z, 3)\}$ (or it might be $\{(x, 0), (y, 1), (z, 2)\}$ if the context tells you our indexing starts at 0).

Note that this definition supports lists of length 0 and 1 (and lists $[x, y]$ of length 2 are not the same as ordered pairs (x, y)).

Traditional names for sets you already know about: The name $\mathbb{Z}$ is traditional for the set of integers (positive and negative whole numbers and zero), $\mathbb{Q}$ for rational numbers, and $\mathbb{R}$ for real numbers. The name $\mathbb{Z}^+$ is traditional for the set of positive integers, $\mathbb{Q}^+$ for positive rational numbers, and $\mathbb{R}^+$ for positive real numbers. The name $\mathbb{N}$ for the set of natural numbers suffers from an ambiguity as whether 0 is a natural number or not. We view $\mathbb{N}$ as referring to $\{n \in \mathbb{Z} : n \geq 0\}$, which includes 0, the set of nonnegative integers, which seems best because we already have a generally accepted name for the positive integers. Notice that the set $\mathbb{N}$ is the set of sizes of finite sets.

4 Informal remarks about cardinality and counting principles

At this point in the class, our next practical application of sets will be in presenting counting principles. Actual definitions of the concepts I introduce here will appear in future lectures but require a bit more work, but I give an informal summary here.

Informal definition: For any set A, $|A|$ is informally defined as the number of elements in A. $|\emptyset| = 0$ of course. Some sets are infinite, and for these, for the moment, we don't define $|A|$.

We will state a formal definition of $|A|$ later.

Disjoint sets: We say that sets A and B are disjoint iff there is no x such that $x \in A$ and $x \in B$. Equivalently, A and B are disjoint iff $A \cap B = \emptyset$. Note that "iff" is an abbreviation for "if and only if" common in mathematical text.

Additive principle: If A and B are finite sets (meaning, $|A|$ and $|B|$ are defined) and $A \cap B = \emptyset$ (in English, A and B are disjoint) then $|A \cup B| = |A| + |B|$.

Additive principle with compensation for overcounting: If A and B are finite sets, $|A \cup B| = |A| + |B| - |A \cap B|$.

Multiplicative principle: If A and B are finite sets, $|A \times B| = |A| \cdot |B|$. Notice that an element of $A \times B$ can be used to represent a process of choosing an element of A, then choosing an element of B.

Notice that $|A \times B| = |B \times A| = |A| \cdot |B|$, because multiplication is commutative, but it is not true that $A \times B = B \times A$ (I'll do an example).

"Exponential principle": If we define A^n as the set of lists $[a_1, \ldots, a_n]$ of n elements taken from the set A, then $|A^n| = |A|^n$. So for example the set of six letter "words" made from the letters A,B,C,D,E, repetitions allowed, pronouncability not a value, which we can write $\{A, B, C, D, E\}^6$, has 5^6 elements.

This is a consequence of repeated application of the multiplicative principle.

A general list counting principle: If we are counting any set of lists $[a_1, a_2, \ldots, a_n]$ where there are k_1 choices for a_1, and given any choice of $a_1, \ldots, a_i$ there are k_{i+1} choices of values for a_{i+1}, the number of lists in the set is $k_1 \cdot k_2 \cdot \ldots k_n = \prod_{i=1}^{n} k_i$.

For example if we are counting words from the alphabet $\{A, B.C, D, E\}$ in which no letters are repeated, the number of such words is $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$, because there are 5 choices for the first letter, 4 for the second (given the first letter, not the same 4 choices in every case), 3 for the third letter (given the previous two choices), and so forth.

5 More stuff

No universal set, so logical relations are not always implemented as set relations:

We prove a theorem (we do not want to give credence to there being a "paradox" here, as people thought during a crisis of foundations at the beginning of the last century).

Let A be a set. Define $\texttt{Russell}(A)$ as $\{x \in A : x \notin x\}$.

We prove that $\texttt{Russell}(A) \notin A$.

We prove this by contradiction. Suppose $\texttt{Russell}(A) \in A$.

Now consider the status of the sentence $\texttt{Russell}(A) \in \texttt{Russell}(A)$.

This expands to $\texttt{Russell}(A) \in \{x \in A : x \notin x\}$

which is equivalent to $\texttt{Russell}(A) \in A$ and $\texttt{Russell}(A) \notin \texttt{Russell}(A)$

which is equivalent to $\texttt{Russell}(A) \notin \texttt{Russell}(A)$ if (as we have assumed) $\texttt{Russell}(A) \in A$ is true.

But this is absurd. So we have shown that $\texttt{Russell}(A) \in A$ cannot be true (and so that $\texttt{Russell}(A) \notin \texttt{Russell}(A)$), which has the more general consequence that there is no set V such that every object x is a member of V.

It follows that there can be no set relation implementing equality, membership or the subset relation, as the domain of any such relation would have to be V, the nonexistent universal set.

It also follows that for no set A can there be a set of all x not belonging to A (a true complement of A). If such a set existed, its union with A would be V. If a working universe U is understood in a particular context, $U \setminus A$ will play the role of the complement of $A \subseteq U$; but it doesn't contain everything that is not in A.

I extend remarks I made earlier. If I ever write $\{x : P[x]\}$, this does not mean the set of *all* x such that $P[x]$, but rather it means $\{x \in U : P[x]\}$ where U is some universal set which can be understood from context. For example, if we are talking about sets of integers, we might write $\{x \in \mathbb{Z} : P[x]\}$ as $\{x : P[x]\}$, particularly if we have to write many set builder notations and want to save typing. Levin (who knows better) defines $\overline{A}$ as the set of all things not in A. We define it as $\{x : x \notin A\}$, that is, as $\{x \in U : x \notin A\}$ where U is a "universe" understood from

context. We discourage this notation, preferring $U - A$. So for example if E is the set of even integers, it is pretty clear that what I mean by $\overline{E}$ should be the set of odd integers, that is $\mathbb{Z} - E$, but there are plenty of things not in E which are also not in this set. In some weird context I might discuss the set of *real numbers* which are not even integers, and $\mathbb{R} - E$ is probably preferable to $\overline{E}$ for this unless it is *very* clear that we are in a discussion of sets of real numbers.

A definition of the ordered pair as a set (historical, easier than the usual one (?)): The first definition of the ordered pair (x, y) given by Norbert Wiener in 1914 (our official one was given by Kuratowski in 1920) was $(x, y) = \{\{\{x\}, \emptyset\}, \{\{y\}\}\}$. If you like solving logic puzzles, you might have fun figuring out why it is easier to extract x and y from this "pair". Hint: this set definitely has two elements, and it has one element with one element and one element with two, whether $x = y$ or not. You are in no way responsible for this. But it is worth noticing that definitions of this kind of concept can take different forms: all we need of the ordered pair is that (x, y) exists for any x and y, and that given (x, y), we can identify its first projection x and its second projection y.

We could define quantifiers using set builder notation: We could define the sentence $(\forall x \in A : P(x))$ as $\{x \in A : P(x)\} = A$. This is read, for all $x \in A$, $P(x)$.

We could define the sentence $(\exists x \in A : P(x))$ as $\{x \in A : P(x)\} \neq \emptyset$. This is read, for some $x \in A$, $P(x)$, or there exists $x \in A$ such that $P(x)$.

4

Defining quantifiers in terms of sets might be taken as an odd maneuver. It is equally odd in discrete math texts (I think odder) that quantifiers are often introduced before sets are introduced...but with set bounds as here, so they depend on informal understanding of sets anyway.

[4]Further, $(\forall x \in A : P(x) \to Q(x))$ is definable as $\{x \in A : P(x)\} \subseteq \{x \in A : Q(x)\}$ ($P \to Q$ can be defined as $(\forall x \in A : P \to Q)$, x not occurring in P, Q and A nonempty). This is not as absurd as it looks: quantified implication was defined first in the actual history! Similar maneuvers can define the other propositional connectives: we content ourselves with observing that $(\forall x : \neg P(x))$ is definable as $\{x \in A : P(x)\} = \emptyset$ and that all the propositional connectives can be defined in terms of negation and implication.

We can then define $\{(x, y) \in A \times B : P(x, y)\}$ as

$$\{u \in A \times B : (\exists x \in A : (\exists y \in B : u = (x, y) \wedge P(x, y)))\}.$$

This can be used as a general model for how to treat complicated expressions appearing left of the colon in set builder notation.

And then we can say in general that a function definition $y = F[x]$ where $F[x]$ stands in for some complicated expression in x is equivalent to

$$f = (A, B, \{(x, F[x]) \in A \times B : x \in A\}),$$

where A is the intended domain (often implicit in a definition of this kind), B is the intended codomain, and the definition only succeeds if for every $x \in A$ it is the case that $F[x] \in B$ (though this can be qualified: in calculus you often work with partial functions, which may be undefined at some elements of the implicitly understood domain: the calculus definition of domain is more analogous to what we call preimage above).

The notation $(x \in A \mapsto F[x] \in B)$ is convenient for this. The mention of the domain A is often omitted, and mentioning the codomain as I do here would be strange in practice.

Reasons why we should only identify functions with their graphs with care:

Suppose we did identify functions (and relations) with their graphs. Two problems would arise for our presentation.

The notion of surjection would not make sense (injection still would). The difficulty is that the codomain of a function cannot be deduced from its graph. Many discrete math books (including ones I have taught from at Boise State) actually define relations and functions as sets of ordered pairs and then try to define onto/surjection as above, which is an error. It is a disgrace that this mistake is so common, I cannot understate this.

One would have to define "f is onto B" as "for any y in B there is x such that $y = f(x)$", for any set B including the range of f (which can be computed from the graph), and define surjection from A to B rather than surjection in any absolute sense.

The above is a general problem and something to watch out for in math texts.

The second problem is particular to our implementation, and is the reason I chose to use explicit domains and codomains here, though by temperament I prefer to identify a function with its graph.

Suppose that G is a set of ordered pairs. How can we show that there are sets A, B such that (A, B, G) is a relation? The quick answer is that we cannot in general with the axioms for set theory which we have presented. We briefly indicate what is needed.

A stronger presentation of the axioms allows a more general construction of unions: given any set A, we can construct the union of all of its elements, which can be written $\bigcup A$, the set of all x such that there is $a \in A$ such that $x \in a$. You can check that $\bigcup\{A, B\}$ is our $A \cup B$: the usual axiomatization of set theory asserts that all unordered pairs $\{x, y\}$ exist and that $\bigcup A$ exists for every set A. This is a much more sophisticated notion of union than the notion of "binary union" which I have used, which should actually already be familiar to you. And understanding the notion of domain and codomain of a relation really does not depend on this more sophisticated notion of union.

This can be codified in a stronger axiom:

Axiom 5b (set union): For any set A, there is a set $\bigcup A$ such that for all x, $x \in \bigcup A$ if and only if there is $a \in A$ such that $x \in A$.

The idea is that $\bigcup A$ is the union of all the sets (possibly infinitely many) which are elements of A: for example, you could check that $\bigcup\{A, B\} = A \cup B$.

Given this concept, the preimage and the image of a relation with graph G an arbitrary set of ordered pairs can be defined as subsets of $\bigcup(\bigcup G)$, so we can present a relation with that graph.

Usually this is not an issue, because before we define a function we know what its intended domain A and codomain B are, so if we are given the rule $y = f(x)$ for the function, we can define $f = (x \mapsto f(x))$ as $(A, B, \{(x, y) \in A \times B : y = f(x)\})$, or, if you do not like the use of (x, y) to the left of the colon [but you should when you see the alternative],

$$(A, B, \{z \in A \times B : (\exists x \in A : \exists y \in B : y = f(x) \wedge z = (x, y))\}).$$

Just for amusement value, the same effect as that of axiom 5b can be achieved by this axiom, which doesn't appear to introduce a mysterious new operation:

Axiom 5c: For any set A all of whose elements are sets, there is a set B such that $A \subseteq \mathcal{P}(B)$.

If you can see how to use axiom 5c to prove the existence of $\bigcup A$ (with the cooperation of axioms 1-5) send me an e-mail and get a bonus point.

If we have the occasion to use the stronger form of the axiom of union, we will mention this explicitly (as the axiom of Set Union) rather than Binary Union, our basic axiom).

A definition of the natural numbers (optional): It is possible to define the natural numbers as sets. It is possible to do this in more than one way (as we defined the ordered pair above in one way, and showed another possibility afterward). We will only give one definition, and really we do this just to show that such a definition is possible. We could equally well assume that natural numbers are individuals.

The definition usually used now is as follows: define 0 as $\emptyset$. Define $\sigma(x)$ (for any set x) as $x \cup \{x\}$.

For any natural number n, the intention is that $\sigma(n) = n + 1$.

So 0 is defined as $\emptyset$, 1 as $\{0\}$, 2 as $\{0, 1\}$, 3 as $\{0, 1, 2\}$, and so forth. Notice that we appear to have each natural number n defined as a set with n elements.

We then provide

Definition: We call a set I *inductive* just in case $0 \in I$ and for any $x \in I$, $\sigma(x) \in I$.

Axiom 6: There is an inductive set.

Definition: We define $\mathbb{N}$ as the set of all objects n such that n belongs to all inductive sets. This implements the usual set of natural numbers. It exists by the axiom of separation and axiom 6: given an inductive set I (by axiom 6), $\mathbb{N}$ is the set of all elements of I which belong to all inductive sets, which exists by separation.

Definition: If n is a natural number and A is a set, and there is a bijection from A to n, we say $|A| = n$. For any set A, we say that A is finite iff there is a natural number n such that $|A| = n$.