What *is* a Model of Set Theory

Lovkush Agarwal

0.1 (Quick) Review of what a model is

You have language $\mathcal{L}$ consisting of function symbols (each with an associated arity), relation symbols (each with an associated arity), and constant symbols, and a theory T, which is a set of sentences (also called axioms), where a sentence is a meaningfully put together string of symbols (symbols can be logical symbols, equality or the symbols in $\mathcal{L}$).

A model $\mathcal{M}$ of theory T consists of a set M (I think it is called the 'carrier set'), and for each $f \in \mathcal{L}$ a function $f^{\mathcal{M}} : M^n \to M$, where n is the arity of f, for each $R \in \mathcal{L}$ a relation $R^{\mathcal{M}} \subset M^n$, where n is the arity of R, for each constant symbol $c \in \mathcal{L}$, an element $c^{\mathcal{M}} \in M$, and these are all chosen so that $\forall \phi \in T, \mathcal{M} \models \phi$.

One example is $\mathcal{L}_G$, the language of groups (so it will have a symbol for the product $(*)$, inverse $(^{-1})$, and identity element e), with sentences consisting of the axioms of group theory (e.g. $\forall x(x * e = e * x = e)$). Then a model of this theory will be a group. Of course, if you take $\mathcal{L}_G$ with the empty theory, then a model would be a set with *any* binary function, unary function, and a specified element (so in particular, it does not have to be a group).

0.2 How to use concept of a model

When given some mathematical structure, there are three stages needed in an attempt to describe it within this framework: choose a language, give meaning to the different aspects of a model, and choose axioms.

For example, you may want to talk about a set of elements which are arranged in a line. So we take $\mathcal{L} = \{R\}$ where R is a binary relation, we give $(x, y) \in R^{\mathcal{M}}$ the meaning 'x is before y in the arrangement' and we take T be the axioms reflexivity, anti-symmetry, transitivity and trichotomy. As you will know, this is exactly the axiomatisation of a total order, and this is the formal concept of a line of elements.

Note that the meaning is 'outside' the model, in the sense that it is not a part of the mechanics of the formal construction of a model or theorems - it is how we as users of the model make sense of the completely formal constructions, i.e., the meanings are how we make the models and theorems meaningful!

0.3 Binary Relations

Let $\mathcal{L} = \{R\}$, where R is a binary relation symbol. Given a model $\mathcal{M}$, we write xRy to mean $(x, y) \in R^{\mathcal{M}}$. For each $x \in M$, let $R_x = \{y \in M : yRx\}$. Notice that specifying a relation R on M is equivalent to specifying R_x for each x. Since R_x can be any subset of M, specifying the relation R is the same as specifying a function, $r : M \to P(M)$ (where P is powerset). This alternative way of thinking about binary relations is conceptually useful.

0.4 And now to...

Collections! I am purposefully using the word 'collection' because 'set' is being used to name the objects used in creating models. So I should have entitled this document 'What is a Model of Collection Theory'. Of course, a set *is* a collection, but keep the concepts seperate for the sake of clarity and understanding.

Now, since the theory of collections is described by the phrase "a is in the collection b", we choose the language $\mathcal{L} = \{R\}$, where R is a binary relation symbol. Next, I will show how a model would describe collections, i.e., I will give meaning to the elements of M, and the relation R:
-every element x of M is (interpreted to be) a *label* of some collection.
-we interpret xRy as 'x is in y'. This is the same as saying we interpret r as '$r(y)$ is the collection of things in y', where r was defined in the previous section. By the previous bullet point, these things are in fact labels.

Some remarks about this choice of meaning:
-Any standard description of collection theory will say that x *is* a collection, as opposed to x *labels* a collection. In my opinion, the latter is the (slightly more) correct way of thinking about it; the reason is subtle, but it is the same reason that the word 'table' is distinct from an actual table. Of course, when doing collection theory, this distinction is irrelevant. (Please excuse this obscure habit of saying collection theory instead of set theory; I am doing it to ensure that we keep the (now formal) concept of a collection distinct from the intuitive notion of set we use when defining a model).

-Another way of looking at the second bullet point is that the function r determines which subsets of M we will consider to be collections (you should see now why I want to keep the notion of set and collection distinct), and which label we would like to assign to those collections.

-Cantor's Theorem (that there is no surjection from A to $P(A)$) has quite a fatal consequence: given any model $\mathcal{M}$ in this language (even before specifying any axioms), there will be some subset of M which will not be a collection. But a subset of M *is* (in the intuitive sense) a collection, so we are already doomed to fail in our attempt at describing collections. (Note that this isn't uncommon in mathematics, e.g. when defining continuity you conclude that some wacky function, that is in no way intuitively continous, is formally continuous. Maybe there is a different framework (i.e. not first order logic) which will allow for formal descriptions which do meet our intuitive notions).

0.5 The Axioms of Collection Theory

Intuively, we would like all subsets of M to be considered as collections, but as I have mentioned, this is not possible. Even so, we would still like our model to be as close to our intuitive notion as possible. One natural idea would be to take the following axiom (actually an axiom scheme, and is known as comprehension): for every formula with one variable $p(x)$,$\{y : p(y)\}$ is a collection. (Formally, we say that for every property p, we have the axiom $\exists x \forall y (yRx \leftrightarrow p(y)))$.

This seems like a perfectly reasonable axiom, and in some sense, you want this to be the case. Unfortunately, Russell's paradox demonsrates that this leads to a contradiction (consider the (1st order!) formula $x \notin x$), and our theory would not have any models.

So we need to modify this axiom, and there are various ways to do this. The most accepted alternative is that for every property p, we have the axiom $\forall a \exists x \forall y (yRx \leftrightarrow (p(y) \wedge y\mathrm{R}a))$, which says that you can form the collection of things which have a certain property AND which are contained in some other collection we already know to be a collection. This resolves the paradox. This however means that we need to specify at least one thing to be a collection, before comprehension can be of any effect. There will be various methods of doing this, and the most widely known are the axioms of ZF (or ZFC). A full discussion of these axioms can be found at `''www.dpmms.cam.ac.uk/~tf/cam_only/axiomsofsettheory.pdf''`

There are other alternative set theories, e.g. NF (New Foundations) restricts comprehension by specifying that the formulas in the comprehension axiom must be stratified. Again, a full description of this alternative is available online.

One final remark I would like to make regards the Axiom of Extensionality. Given how I have interpreted the model (elements are labels of collections), there is strictly no need for the Axiom, as the axiom would simply state that a collection has a unique label, which is equivalent to saying that r is injective. Having multiple labels for the same collection has no consequence (I do not think) to the underlying structures and relationships between the collections, as you could take the quotient of a model M without extensionality, using some clever equivalence relation, to get a model M', which has extensionality.(Note, I do not know if this is correct or not, it is just a hunch.)

Another final remark is that no model of collection theory is actually known; so it is possible that there is some contradiction in ZF which we simply have not found yet. Also, if you think the model has to be very big and scary, you'd be pleasantly surprised; if any model did exist, a countable model (say with carrier set $\mathbb{N}$) would exist (via Downward Lowenheim-Skolem), so the model would only have to be scary. What's particularly crazy about this is that from the axioms of ZF, you can conclude that collections of arbitrary cardinality must be in the model, yet, all subsets of $\mathbb{N}$ are by definition countable. And no, this is not a contradiction, but I will let you ponder about this.