Models of Set Theory

You may have noticed that, when introducing the axioms of ZFC, we never *really* answered the question "*What is a set?*" Instead, we developed a formal theory of axioms for a binary relation that somehow describe "*how sets work*", that is, how we can obtain sets from given ones using well-known operations like power set and union.

We have then seen how we can develop a lot of standard mathematical *objects* (like $\mathbb{N}$, $\mathbb{R}$) and *techniques* (like induction and definition by recursion) **inside** this formal system. In fact, most of mathematics can be developed formally inside this system. Almost all proofs you find in any standard math book are proofs that can be formalized in ZFC. It is very tedious to do this for us humans, but there is little doubt it can be done, and in fact, looking at the recent work on **proof assistants** (like Coq or Lean), many parts of mathematics have been formalized (albeit not directly in ZFC).

This expressiveness gives ZFC its foundational importance, but it is also the cause for much confusion for someone who first studies set theory.

From a pedagogical point of view, in what follows it is helpful to assume a "**Platonist**" perspective of mathematics, and set theory in particular, namely that *sets and the relations between exist independently (and outside) of the ZFC axioms*. The set of real numbers exists, and our development of $\mathbb{R}$ inside ZFC is just a formal way to describe them. From this perspective, the axioms of ZF (AC is a little different) are just obvious truths about sets, just like the **Peano axioms** are obvious truths about natural numbers.

Among other things, this perspective allows us to treat ZF just like any other mathematical theory, like *group theory* or the theory of *algebraically closed fields*. In particular, we can think about **models of set theory** the way we would think about models of group theory, in the sense of model theory.

A **model** would simply be a set M together with a binary relation E on S such that

$$(M, E) \models \mathsf{ZF}, \tag{1}$$

that is, all axioms hold when interpreted in (M, E). Note that we use "*set*" in this context not in the formal sense, but in the "meta"-sense (the Platonist world of sets).

Working in the meta-theory ("*that what is mathematically true*"), we know by Gödel's completeness theorem that

> if ZF is consistent, then it has a model.

This model should be seen as a **set-theoretic universe**: Its elements can be seen as sets, and the interpretation E of the $\in$-symbol will tell us how these sets are connected via the element-relation.

Note that E does not have to be the *actual* element relation on a set (of sets), but just some binary relation so that the axioms are satisfied.

In the meta-world, there are, of course, sets other than M, but that does not matter here, since all we are interested in is giving *some* universe in which our axioms hold. (Timothy Chow has suggested that set theory should rather be called "*universe theory*. He is right in the sense that what axiomatic set theory does is to define such *universes of sets*, rather than what a set is.)

In the meta-theory, we can then follow the usual techniques to show provability or non-provability results.

If we want to prove that CH is consistent with ZF (assuming ZF is consistent), we need to find a model in which both hold.

One difficulty in working with models of set theory is that they can look very different depending on whether you look at a model "**from the inside**" or "**from the outside**".

To illustrate this, assume ZF is consistent. Then, by the **Löwenheim-Skolem theorem**, there exists a **countable model** for ZF. Yet it is a theorem of ZF that *there exists an uncountable set*. This is often referred to as **Skolem's paradox**, although it is not really an antinomy.

If we break this down a bit, we see that the apparent paradox is really just a matter of perspective (*inside* or *outside*). Assume (M, E) is a countable model of ZF. Then there exists $x \in M$ such that there is no injection from x to the natural numbers. Since M is countable, x can have at most countably many elements. So why is this *not* a contradiction? We should really read the statement above as

there is no injection **in** M from x to M**'s version** of the natural numbers.

In other words, even though x is countable *from the outside*, x appears uncountable *inside* M since a mapping witnessing its countability does not exist in M.

This is a first warning sign that models of ZF can behave in very unexpected ways. For another example, recall the axiom of *Foundation* asserts that the $\in$-relation is well-founded. But again, this means only "from the inside".

Proposition 0.1. *If* ZF *is consistent, than there exists a model* (M, E) *of* ZF *such that* (M, E) *is ill-founded.*

Proof. Introduce new constant symbols c_n ($n \in \mathbb{N}$) and add the formulas $\varphi_n \equiv c_{n+1} \in c_n$ to the axioms of ZF. It is not hard to show, using the compactness theorem, that $\text{ZF} + \bigcup_n \varphi_n$ has a model (M^*, E^*), for which the set $\{c_n : n \in \mathbb{N}\}$ is ill-founded.

$\square$

Since, as mentioned above, the model (M^*, E^*) satisfies *Foundation*, the set $\{c_n : n \in \mathbb{N}\}$ is actually *not in* the model (and neither can be any other set with an infinite descending $\in$-chain).

1　Mostowski collapse

If we restrict ourselves to models on which the E-relation is *actually* well-founded (i.e. from the outside), then interestingly these models look, in a way, "*natural*": They can be assumed to be the $\in$-relation on a set. Such models are also called **standard**.

Given a set theoretic structure (M, E) (not necessarily a model of ZF), for each $x \in M$ let

$$\text{ext}_E(x) = \{y \in X : y \, E \, x\}$$

If E behaves "set-like", then it will respect the axiom of *Extensionality*, i.e. two sets are identical if and only if they have the same elements. Therefore we say that E is **extensional** if

$$x, z \in X, \ x \neq z \quad \text{implies} \quad \text{ext}_E(x) \neq \text{ext}_E(z).$$

Furthermore, as stated above, we want to exclude infinite descending E-chains. Recall we say that E is **well-founded** if

every non-empty set $Y \subseteq X$ has an E-minimal element.

Theorem 1.1 (Mostowski collapse)**.** *If E is an extensional and well-founded relation on a set X, then there exists a transitive set S and a bijection $\pi : X \to S$ such that*

$$x \, E \, y \iff \pi(x) \in \pi(y) \quad \text{for all } x, y \in X.$$

Moreover, S and π are unique.

Proof. We construct π and $S = \mathrm{im}(\pi)$ by recursion on E, which is possible since it is well-founded.

For each $x \in X$, let
$$\pi(x) = \{\pi(y) \colon y \, E \, x\},$$
and set $S = \mathrm{ran}(\pi)$.

The injectivity of π follows from the extensionality of π by induction along E: Suppose we have shown
$$\forall z \ (zEx \rightarrow \forall y \in X(\pi(z) = \pi(y) \rightarrow z = y)).$$
and we have to show that it holds for x. Assume $\pi(x) = \pi(y)$ for some $y \in X$. Then

$$
\begin{aligned}
cEx \Rightarrow & & \pi(c) \in \pi(x) &= \pi(y) & \\
\Rightarrow & & \pi(c) &= \pi(z) & \text{for some } zEy \\
\Rightarrow & & c &= z & \text{(by ind. hyp., since } cEx) \\
\Rightarrow & & cEy. & &
\end{aligned}
$$

Similarly, we get $cEy \Rightarrow cEx$, hence $x = y$ as desired due to the extensionality of E. Finally we have

$$
\begin{aligned}
\pi(x) \in \pi(y) \Rightarrow & & \pi(x) &= \pi(c) & \text{for some } cEy \\
\Rightarrow & & x &= c & \text{(since } \pi \text{ is injective)} \\
\Rightarrow & & xEy & &
\end{aligned}
$$

Thus π is an isomorphism.

To see the uniqueness of π and S, assume ρ, T are such that the statement of the theorem is satisfied. Then $\pi \circ \rho^{-1}$ is an isomorphism between $(T, \in)$ and $(S, \in)$. Now apply the following lemma.

$\square$

Lemma 1.2. *Suppose X, Y are sets, and θ is an isomorphism between $(X, \in)$ and $(Y, \in)$. Then $X = Y$ and $\theta(x) = x$ for all $x \in X$.*

Proof. By induction on the well-founded relation $\in$. Assume that $\theta(z) = z$ for all $z \in x$ and let $y = \theta(x)$.

We have $x \subseteq y$ because if $z \in x$, then $z = \theta(z) \in \theta(x) = y$.

We also have $y \subseteq x$: Let $t \in y$. Since $y \in Y$, there is $z \in X$ with $\theta(z) = t$. Since $\theta(z) \in y$ and $y = \theta(x)$, we have $z \in x$, and thus $t = \theta(z) = z \in x$.

Hence $x = y$, and this also implies $\theta(x) = x$.

$\square$

Corollary 1.2.1. *If E is a well-order on a set X, then (X, E) is isomorphic to $(\alpha, \in)$ for a unique ordinal α.*

Proof. Every linear order is extensional. Hence we can apply the Mostowski Collapse. It is easy to see that the resulting set is an ordinal.

$\square$

2 Absoluteness and transitive models

Even if we consider well-founded standard models, interpreting set-theoretic statements in them can still lead to very different results, even for very simple formulas.

Example 2.1. *Consider $M = \{0, \{\{0\}\}\}$ and let $\varphi \equiv \{\{0\}\} \subseteq 0$. Clearly φ is not true from the "outside". But what if we look at it from "inside" M? $x \subseteq y$ means $\forall z(z \in x \rightarrow z \in y)$. Inside M this means every set z in M that is an element of x is also an element of y. But in M, there is no element $z \in \{\{0\}\}$, $\{0\}$ is not an element of M. Therefore,*
$$(M, \in) \models \{\{0\}\} \subseteq 0. \tag{2}$$

In the example above, the set M is not transitive, which allowed it to "hide" its elements. It turns out that if we require our model to be transitive, the truth of simple formulas cannot vary between the "inside" and the "outside" perspective. We call such properties **absolute**.

Given a formula φ and some class M, we can **relativize** φ with respect to M essentially by restricting all quantifiers in φ to range over M, i.e. $\exists x$ becomes $\exists x \in M$ and $\forall x$ becomes $\forall x \in M$. Note that classes are defined by formulas, so the resulting formula, which we denote by φ^M, is still a formula of set theory.

Definition 2.1. *We say a formula $\varphi(x_0, \ldots, x_n)$ is **absolute** for a class M if for all $a_0, \ldots, a_n \in M$*

$$\varphi^M(a_0, \ldots, a_n) \text{ holds} \iff \varphi(a_0, \ldots, a_n) \text{ holds.} \tag{3}$$

Definition 2.2. *A formula φ is Δ_0 if it has no quantifiers or every quantifier of φ is of the form $\exists x \in y$ or $\forall x \in y$ (i.e. every quantifier is bounded).*

Lemma 2.1. *If M is transitive and φ is Δ_0, then φ is absolute for M.*

Sketch. Clearly $x = y$ and $x \in y$ are absolute for any M. It is also not hard to see that if φ and ψ are absolute for M, then so are $\neg\varphi$ and $\varphi \wedge \psi$. Hence all quantifier free formulas are absolute.

Finally, if φ is absolute for M, so is $\psi \equiv \exists x \in y\, \varphi$: If $\psi^M(y, \bar{z})$ holds for $y, \bar{z} \in M$, then we have $[\exists x(x \in y \wedge \varphi(x, y, \bar{z}))]^M$, i.e., $\exists x \in M(x \in y \wedge \varphi^M(x, y, \bar{z}))$. Since $\varphi^M(x, y, \bar{z})$ if and only if $\varphi(x, y, \bar{z})$, it follows that $\exists x \in y\varphi(x, y, \bar{z})$, i.e. ψ.

Conversely, if for $y, \bar{z} \in M$, $\exists x \in y\varphi(x, y, \bar{z})$, then since M is transitive, x belongs to M, and since $\varphi(x, y, \bar{z})$ if and only $\varphi^M(x, y, \bar{z})$, we have $\exists x \in M\,(x \in y \wedge \varphi^M(x, y, \bar{z}))$ and so $\psi^M(y, \bar{z})$.

$\square$

Proposition 2.1. *The following expressions can be written as Δ_0-formulas and hence are absolute for any transitive class M:*

- $x = \emptyset$, $x = y$, $x \subseteq y$, $z = x \cap y$, $z = x \cup y$, $x = \bigcap y$, $x = \bigcup y$, $z = \{x, y\}$, $z = (x, y)$, $z = x \times y$

- r *is a relation,* f *is a function,* $x = \mathrm{dom}(f)$, $y = \mathrm{ran}(f)$

- $<$ *is a linear order on* x, x *is an ordinal,* x *is a limit ordinal,* $x = \omega$

We leave the proof as an exercise. We can put this absoluteness property to use and identify many "almost" models of ZF.

Theorem 2.2. *If $\alpha > \omega$ is a limit ordinal, then $(V_\alpha, \in) \models \mathsf{ZFC} - \mathsf{Replacement}$.*

Sketch. We verify a few axioms and leave the rest as an exercise.

Extensionality: The relativized version of this axiom is

$$\forall a, b \in V_\alpha\, (\forall x \in V_\alpha(x \in a \leftrightarrow x \in b) \to a = b) \tag{4}$$

This simply states that the $\in$-relation (as a binary relation) is extensional on V_α (as defined above). It is easy to see that for every transitive set, the $\in$-relation is extensional.

Power Set: We have seen that $\subseteq$ is absolute for transitive classes. Therefore, the relativized version of the Power Set axiom becomes

$$\forall a \in V_\alpha\, \exists y \in V_\alpha \forall z \in V_\alpha(z \in y \leftrightarrow z \subseteq a) \tag{5}$$

This means we only need to have those subsets in the power set that are in V_α. In other words, we need to verify

$$\forall a \in V_\alpha\, (\mathcal{P}(a) \cap V_\alpha \in V_\alpha) \tag{6}$$

If $a \in V_\alpha$, then $a \in V_\beta$ for some $\beta < \alpha$, since α is limit. Since V_β is transitive, $a \subseteq V_\beta$ and thus $\mathcal{P}(a) \subseteq \mathcal{P}(V_\beta) = V_{\beta+1}$, which in turn yields $\mathcal{P}(a) \in V_{\beta+2}$

If $b \in \mathcal{P}(a) \cap V_\alpha$, then $b \in V_{\beta+2}$. Therefore, $\mathcal{P}(a) \cap V_\alpha \subseteq V_{\beta+2}$. It follows that $\mathcal{P}(a) \cap V_\alpha \in V_{\beta+3} \subseteq V_\alpha$, as desired.

Infinity: We have seen that $x = \omega$ is absolute for transitive classes. Hence V_α will satisfy the Axiom of Infinity if $\omega \in V_\alpha$. But this is true since we assume $\alpha > \omega$.

$\square$

What is the problem with Replacement? The axiom says that if φ defines a function, then the image of any set under this function is a set. Relativized to some M, this means we need to find a set *in M* that contains the image. Here we run into a cardinality problem. The cardinality of V_α can be much bigger than the cardinality of α. For example, $\mathcal{P}(\omega) \in V_{\omega+\omega}$ but $\omega + \omega$ is countable. We could have a function that maps $\mathcal{P}(\omega)$ to sets of rank cofinal in $\omega + \omega$, which implies that the image cannot be in $V_{\omega+\omega}$.

For V_α to be a model, we would need α to be "unreachable" by such mappings. This gives rise to the notion of *inaccessible* cardinals, which we will introduce in the next section.