BEYOND INFINITY 6

JADE NINE

1. Models

A model is a collection of objects which satisfies a theory (remember, a theory is a collection of axioms and logic). The point of a model is to give a theory enough playing blocks and blueprints that it can satisfy all needs it possibly has. In set theory, a model is just a collection of sets where every set that an axiom says should exist, does exist.

Here is a simple theory, which I will call the *Theory of Steps*:

Axiom 1 (Existence)**.** $\exists x$

Axiom 2 (Step up)**.** $\forall x(\exists x^+(x^+ \neq x))$

Axiom 3 (Step down)**.** $\forall x(\exists x^-(x^- \neq x))$

Axiom 4 (Navigation)**.** $\forall x((x^+)^- = (x^-)^+ = x))$

Can you think of a model for this theory?

A set we have introduced in this course actually fulfills the requirements of a model for the Theory of Steps. The integers, $\mathbb{Z}$, with the operations $+, -$ defined as $+1, -1$, satisfies all the axioms: there exists an integer (the set is not empty), for each integer there is a $x^+ = x + 1$ and a $x^- = x - 1$, and the actions + and - are inverses of each other (they cancel each other out). Draw out a diagram for this model, and use the axioms to move around between the various points in the model. Notice that, no matter which axiom you pick, the destination stipulated by the axiom is available in the integers - for example, if you're at 5 and you want to move to 5^+, you can move to $5 + 1 = 6$.

Can you think of any other models?

In fact, two disjoint (separate) copies of the integers also satisfies the theory. Every axiom is still true if you have two disjoint copies of the integers. Points in the first copy won't be able to reach points in the second copy by stepping up or down, but for all the theory cares about, all the steps it could want to take are available.

You can think of other models for this theory as well. $\mathbb{Z} \times \mathbb{Z}$ is another model, where the $+$ and $-$ actions move along one particular direction - for example, (m, n) can move to $(m + 1, n)$ or $(m - 1, n)$. The rationals, $\mathbb{Q}$, are also a model: every rational number q has both a $q + 1$ and a $q - 1$ available; there may be many numbers between q and $q + 1$, but that doesn't violate any rules.

So let's make a hypothesis:

Axiom 5 (Interstitial Hypothesis)**.** $\forall x(\neg\exists y(x < y < x^+))$

To be perfectly clear about this, I also need an axiom that states that $x^- < x < x^+$ as well as axioms that assert that $<$ is a partial order relation (irreflexivity, antisymmetry, transitivity). But if I do that, then the Interstitial Hypothesis makes sense. Is it true? Is it false?

The integers, $\mathbb{Z}$, are a model for the Theory of Steps in which the Interstitial Hypothesis is true. This is our first important result concerning models.

Theorem 1 (Model Satisfaction)**.** *Given a theory T and a statement p, if there is a model M for T in which p is true, then p cannot be disproved by the axioms and logic of T.*

Therefore, the Theory of Steps cannot prove that the Interstitial Hypothesis is false. If it could, then the integers $\mathbb{Z}$ would not be a model for the theory, but they are.

Is the Interstitial Hypothesis true, then? Well, the Interstitial Hypothesis cannot be proven to be true in the Theory of Steps, because it is false in the model $\mathbb{Q}$. Specifically, $\mathbb{Q}$ is a model of the Theory of Steps in which the Interstitial Hypothesis is false, therefore the hypothesis cannot be proven in the Theory of Steps.

Date: August 2, 2020.

This can be restated as follows - this is equivalent to the Model Satisfaction theorem above, one simply takes the statement $\neg p$ instead of p in order to turn one formulation into the other.

Theorem 2 (Model Satisfaction)**.** *Given a theory T and a statement p, if there is a model M for T in which p is false, then p cannot be proved by the axioms and logic of T.*

If a statement p cannot be proved in a theory T, then the extended theory T+$\neg p$ is consistent - in other words, saying that T is unable to prove p is equivalent to saying that T is not able to derive a contradiction from "$\neg p$".

This leads us to the last important concept for this section:

Definition 1 (Independence)**.** Given a theory T and a statement p, if the extended theories T+p and T+$\neg p$ are both consistent (assuming T is consistent), then p is independent of T.

Recall that, in the previous sheet, we described independence as the notion that a statement can be neither proved nor disproved. This definition simply says that an equivalent conceptualization of independence is that the statement and its negation are both consistent with the theory.

2. Constructible Model

The simplest model for a theory is the constructible model of the theory, which is a model that you build up by starting with some basic notion of 0 or emptiness and then adding in every object that the axioms prove must exist because of a previously introduced object.

In the Theory of Steps, the constructible model comes from starting at 0, and then repeatedly adding in points that you can reach *from points already added.* So we start with $\{0\}$, then we get $\{-1, 0, 1\}$, then we get $\{-2, -1, 0, 1, 2\}$, etc. At each step, we use the fact that, for any object x, the existence of x implies the existence of itself (thus x itself gets to stay in the model) as well as the existence of x^+ and x^-.

The integers, therefore, are the constructible model of the Theory of Steps. In this constructible model, the Interstitial Hypothesis is true.

It is only when we move to a 'bigger' model with more objects that we see that the Interstitial Hypothesis is false.

The constructible model of a theory is always a good place to start if one wants to prove independence results in the theory.

3. Continuum Hypothesis

The Continuum Hypothesis, or CH, is fascinating - it concerns the question of whether there are cardinal numbers between ω and $\mathfrak{c}$, and it turns out to be independent of the (generally more intuitive) axioms of the theory ZFC - so we don't actually know if such cardinal numbers exist, unless we are willing to expand our intuition past ZFC. If CH is true, then every ordinal in $\mathfrak{c}$ is countable, i.e. every subset of $\mathbb{R}$ is either finite, countable, or bijectable to all of $\mathbb{R}$; this notion allows some curious deductions. Alternatively, assuming that every ω_1-size subset of $\mathbb{R}$ is a proper subset of $\mathbb{R}$ can also be used to deduce some interesting notions. It is easier to appreciate the implications of CH and $\neg$CH when one studies the real number line, and various mathematical theories that concern it (analysis, measure theory, etc.), but for the moment, the fact alone that there are numbers whose existence we don't know and can't be confirmed, is itself intriguing.

To prove that CH is independent of ZFC, we will need a model in which it's true and a model in which it's false. The issue is that ZFC, unlike the Theory of Steps, is far more complex, and we can't obviously look at a set and say 'yes, that satisfies all our axioms'; we will need some tools to build the models we need for this theorem. Recall that, because of the Axiom of Separation, any logical property of sets is relevant to ZFC, and there are so many such possible properties that there is no easy, simple set that we can look at and know that all the ZFC-buildable sets exist inside it.

To get the models we need, we will assemble models that are put together in a way that guarantees all the axioms of ZFC are satisfied.

The first model is the constructible model of ZFC. It is more difficult to assemble than the constructible model of the Theory of Steps, but the core idea is the same: start with the empty set, and then add all sets that can be proved to exist by applying ZFC axioms to the empty set. Then, repeat this process - we will need to repeat this process for *all ordinals*, in order to ensure that we reach everything. The resulting model is known as the constructible universe, typically written with the capital letter L, and was first described by Gödel. One can provide a formal, exact definition, which puts together the universe L using transfinite induction, but I won't include that here, since I won't formally prove its relevance to CH, I am more interested in you developing the intuition of the relationship.

The second model we will need is built using a powerful, more advanced technique, known as forcing. Forcing is the concept that we can take a model and extend it into a bigger model by rewriting the set-element relation $\in$ with subscripts as elements p from a poset P; then we take a subset $G \subset P$, and declare $x \in_P y$ if $\exists p \in G(x \in_p y)$. Given a model M, a poset P, and a subset $G \subset P$, if we rewrite the sets of M in this way, then the result is called $M[G]$, an extension of the model M. The rewriting of all sets M using $\in_p$ for elements $p \in P$ instead of the plain usual $\in$ is known as the construction of P-names. One can provide a formal, exact definition, which puts together the P-names using a recursive definition; I won't include that here, since we won't formally prove that the forcing technique works: I am more interested in you learning what it means and what it can be used for. For now, know that both the axioms of ZFC hold true in $M[G]$, and the set G itself is in $M[G]$, which is why this technique is useful (we want to be able to add particularly useful sets G to a model).

Why do these models help us?

In the Constructible Universe model, CH is true. In the step that we end up adding ω_1 to L, the ordinal $\mathfrak{c}$, which is a well-ordering of the powerset $\mathcal{P}(\omega)$, can consist only of ordinals in ω_1 (which are all countable), thus the step of adding the powerset of ω assembles a set whose only elements are ordinals whose cardinality is the same as ω. The set of all such ordinals is ω_1, thus the cardinal $\mathfrak{c}$ is the same as the cardinal ω_1 in the Constructible Universe model. This is the general idea.

In the forcing extension model that we will end up using, we start with the constructible universe L and then add in ω_2 many subsets of ω in a way that is consistent with all the axioms of ZFC. Since L itself satisfies all the axioms of ZFC, the extension of L by adding ω_2 many subsets of ω will also be a model of ZFC. This produces a model of ZFC in which CH is false - in particular, we will end up with a model in which $\mathfrak{c} = \omega_2$. (You can pick a larger cardinal instead of ω_2 if you want, but the exact choice is not so important.)

4. Forcing

In order for a forcing extension model to satisfy all of ZFC, the subset $G \subset P$ needs to satisfy certain properties that line up with necessary logical requirements. Specifically, the key notions are transitivity and combination, intuitively similar to the logical notions $\implies$ and $\wedge$; I won't formalize the details of exactly why, but G needs to satisfy the following two properties:

- $\forall p \in G(\forall q > p(q \in G))$. If p is in G, then so is everything bigger than p.
- $\forall p, q \in G(\exists r \in G(r < p \wedge r < q))$. If two elements p, q are in G, then they have a common root, an element r in G which is smaller than both of them. In other words, every pair of elements in G is compatible.

A subset of a poset satisfying these two properties is called a filter. We also need G to be maximal if we want G to end up being a set in $M[G]$ - otherwise, evaluating $\in_p$ for $p \in G$ does not necessarily produce a set that is equal to G. A maximal filter is called an ultrafilter. For now, don't worry about this; it is a concept that is not crucial to the core of the theory.

A subset of a poset satisfying this property: $\forall p \in P(\exists q \in D(q < p))$ is called dense. **This is different than the notion of density discussed earlier**, e.g. the idea that the rationals are a dense poset is not at all related to this definition of dense. It's an entirely separate notion that happens to have the same name.

A filter that intersects every dense set of a poset is called generic.

Given a model M and a poset $P \in M$, if a filter $G \subset P$ intersects every dense set $D \in M(D \subset P)$ (i.e. every dense subset of P that is also an element of M), then G is said to be M-generic. In essence, G hits all the relevant information carried by M that we need related to a particular poset P. There is no need for G to intersect dense sets $D \subset P$ that are not elements of the model M; G need only intersect dense subsets of P that are found inside the model M, thus the name M-generic.

5. Cohen Reals

A Cohen Real (also called Cohen Generic Real), first described by Paul Cohen, is a subset of ω that is added to a model M via forcing.

First, we will extend a model M by adding a single Cohen Real to the model. This serves as an illustration of what I'm talking about, and also should put together all the pieces of why we're using posets. What good is a poset for extending a model? Well, it turns out, an M-generic filter of a certain poset of functions is a Cohen Real.

Theorem 3 (Cohen Real)**.** *Let M be a model of ZFC and let $P = Fin(\omega, 2)$, that is the set of all partial functions whose domain is a finite subset of ω and whose range is 2; the order relation is given by function extension, which I described last week. Then an M-generic filter of P is a subset of ω.*

Proof. Specifically, the M-generic filter G of P will be an entire function from ω to 2. Here's why: for every $n \in \omega$, the set D_n of partial functions in $Fin(\omega, 2)$ whose domain includes n is a dense subset of P. This is true because, given any partial function $f : \omega \to 2$, either $n \in dom(f)$ and therefore $f \in D_n$, or $n \notin dom(f)$, and we can thus extend f to a function f' whose domain is $dom(f) \cup \{n\}$, by setting $f' = f$ on $dom(f)$ and $f'(n) =$(0 or 1).

Thus, an M-generic filter G of P intersects every subset D_n, therefore it contains an element in every D_n. But because G is a filter, all its elements are compatible; note that two partial functions that send the same number to a different value are incompatible, for example the partial functions $h(0) = 0$ and $j(0) = 1$ are incompatible because there is no function that extends both of these (a function cannot map 0 to both 0 and 1 simultaneously). Thus all functions in G agree everywhere they are defined. Therefore, the union of all the functions in G is a single function that agrees everywhere, and is also defined everywhere because it is an extension of a function in every D_n.

Therefore, $g = \bigcup G$ is an entire function from ω to 2. This function can then be interpreted as a subset of ω by collecting $s_g = \{n \in \omega \big| g(n) = 1\}$. Then $s_g \subset \omega$. □

The next step is to add *two* Cohen Reals *at the same time.*

Theorem 4 (Two Cohen Reals)**.** *Let M be a model of ZFC and let $P = Fin(2 \times \omega, 2)$. Then an M-generic filter of P is a pair of distinct subsets of ω.*

Proof. The M-generic filter G will be an entire function from $2 \times \omega$ to 2, for reasons similar to that in the previous proof. However, there is now an entirely new class of dense subsets:

Let $I = \{f \in Fin(2 \times \omega, 2) \big| \exists n(f(0, n) \neq f(1, n))\}$. In other words, I is the set of all functions where $(0, n)$ and $(1, n)$ map to different results for some n. Then I is dense, because any partial function $f \in Fin(2 \times \omega, 2)$ can be extended to a function in I by picking some element $n \in \omega$ that does not appear as the second coordinate of any domain element in f, and then let $f(0, n) =$(0 or 1) and $f(1, n)$ be the other choice (1 or 0).

Therefore, the M-generic filter G is an entire function from $2 \times \omega$ to ω where the $(0, n)$ and $(1, n)$ values disagree for at least one $n \in \omega$.

Now, let $g = \bigcup G$ and interpret g as two functions, g_0 and g_1, defined as $g_0(n) = g(0, n)$ and $g_1(n) = g(1, n)$. Then g_0 and g_1 define two distinct subsets s_{g_0}, s_{g_1} of ω, thus s_{g_0} and s_{g_1} are two distinct Cohen Reals. □

We are now ready to prove the independence of the Continuum Hypothesis.

6. Independence of CH

Theorem 5 (Independence of CH)**.** *The Continuum Hypothesis, $\omega_1 = \mathfrak{c}$, is independent of ZFC.*

Proof. To prove that ZFC+CH is consistent, we must produce a model of ZFC in which CH holds true. The Constructible Universe model L serves this purpose.

To prove that ZFC+¬CH is consistent, we must produce a model of ZFC in which CH is false. The forcing extension L$[G]$ where G is an L-generic filter of the poset $Fin(\omega_2 \times \omega, \omega)$ is a model in which $\omega_2 = \mathfrak{c}$. Thus CH is false in L$[G]$. (Specifically, G adds ω_2 distinct Cohen Reals to L, thus L has its ω_1 many original subsets of ω plus ω_2 many more.) □

There is a catch with this proof: cardinal collapse.

7. Cardinal Collapse

Theorem 6 (Collapse of ω_1)**.** *Let M be a model of ZFC. Let $P = Fin(\omega, \omega_1)$. Let G be an M-generic filter of P. Then $\bigcup G$ is a function from ω onto ω_1, whose domain is all of ω and whose range is all of ω_1.*

Proof. We already know that the sets D_n for $n \in \omega$ of functions whose domain includes n is a collection of dense subsets of P. Then, the sets R_α of functions whose range includes $\alpha \in \omega_1$ is also a collection of dense subsets: indeed, for any finite partial function $f : \omega \to \omega_1$, either α is in the range of f, or f can be extended to a function f' whose range includes α by selecting some $n \notin dom(f)$ and setting $f'(n) = \alpha$. □

Why is this a problem? Because smaller cardinals cannot be domains of functions whose range covers an entire larger cardinal (such a function can easily be modified into a bijection).

So what happened here?

It turns out that generic forcing extensions of models can reveal some underlying truths of set theory that have eluded us so far. I defined uncountable sets for you, and I introduced you to $\mathfrak{c}$ and ω_1 which are both uncountable ordinals, and I proved that each one is uncountable (using proofs by contradiction). However, as it turns out, the theory of uncountable sets does not produce uncountable sets by actually establishing the exsitence of uncountably many distinct elements: rather, it simply asserts that a set exists with a property, and those properties are contradictory with countability.

When I described forcing extensions, I told you that, if all ZFC axioms hold in a model M, they will also hold in $M[G]$. But I did not say that all sets will end up in the same place, and that all properties of sets would remain the same. In fact, some properties of sets will change in the forcing extension. Such properties are called non-absolute. Any property that never changes in a forcing extension is absolute.

Here are some absolute properties:

- $x = \emptyset$
- x is a function
- x is a poset
- x is a well-ordered poset
- x is an ordinal
- $x = \omega$

But the property $x = \omega_1$, or 'x is uncountable', is not absolute. Neither is 'x is countable'.

Remember that we constructed ω_1 by using the axiom of separation on the property 'x is countable'. This property, which is not absolute, will produce a *different* set in the extension $M[G]$ if you try to use it in the axiom of separation. Therefore, in the example $M[G]$ above, the set $\omega_1 \in M$ becomes a countable set in $M[G]$ and the set $\omega_1 \in M[G]$ is some different set in $M[G]$, constructed via the axiom of separation using the non-absolute property of countability. For this reason, for non-absolute objects like ω_1, it is common to write ω_1^M to indicate the set that is constructed via the axiom of separation inside M; for different models M, N, the sets ω_1^M, ω_1^N may be different.

A set being a powerset of another set is also not absolute, and in fact you can collapse $\mathfrak{c}$ onto a countable ordinal also.

Now let's return to the proof of the independence of CH. Note that, for the forcing extension model, I used $Fin(\omega_2 \times \omega, \omega)$. This uses ω_2^L, which is a non-absolute set. In order for the proof to be complete, we must ensure that ω_2^L *does not collapse*, and therefore $\omega_2^L = \omega_2^{L[G]}$. It turns out that, if one imposes a certain condition on the poset P, then cardinals will never collapse. This condition is called the countable chain condition. I will not describe it here, it's enough to know that it exists and that $Fin(\omega_2 \times \omega, 2)$ has this property.

8. Countable Models

Given a model M of ZFC, if the model M is countable, then one can prove that every poset $P \in M$ has an M-generic filter G_P^M.

Why should there be countable models of ZFC?

A theorem establishing the existence of countable models of ZFC was first proved by Löwenheim and Skolem, and is known as the (downward) Löwenheim–Skolem Theorem:

Theorem 7 (Löwenheim–Skolem (downward))**.** *Given any model M of ZFC, if M is uncountable (i.e. contains uncountably many sets), then there is a sub-model $N \subset M$ with smaller cardinality that is also a model of ZFC.*

As long as models of ZFC exist at all - which they will if ZFC is consistent, which is just an assumption of faith we have to make - then we can shrink the model repeatedly until we eventually end up at a countable model.

As you have seen, uncountable sets can collapse into countable sets in model extensions. The theory ZFC itself is countable in a fundamental sense: there are only countably many sentences that can be written in ZFC, because ZFC has finitely many symbols to work with (countably many if you want to include countably many variable names to use), and every statement in ZFC is a finite assembly of these symbols. Thus, there are only finitely many statements. Therefore, it is possible to satisfy every single statement with an assembly of finitely or countably many sets for each statement, the total product of which will be a countable model. Thus ω_1 may be interpeted as having uncountably many elements, but only countably many of them can be distinguished from each other with explicit properties and statements.

9. Beyond ZFC

Here are a few more statements that are independent of ZFC:

Axiom 1 (Suslin). *Given a poset, let* antichain *refer to subsets of incomparable elements in the poset. Then, any tree of height ω_1 has either an uncountable branch or an uncountable antichain. (A tree of height ω_1 whose chains and antichains are all countable is called a Suslin tree.)*

Axiom 2 (Martin). *Given any model M of ZFC and a poset P such that M contains less than $\mathfrak{c}$ many dense subsets of P, there is an M-generic filter G of P.*

Axiom 3 (Inaccessible Cardinal). *There is a cardinal λ such that $\forall\alpha \in \lambda(|\mathcal{P}(\alpha)| \in \lambda)$, and for any ordinal $\gamma \in \lambda$, every subset $g \subset \lambda$ of order type γ is 'smaller' than λ in the sense that $\bigcup g$ is a proper subset of λ.*

Axiom 4 (Zero Sharp). *There is a set, called $0^{\#}$, that encodes all true statements about (a certain collection of) sets in the Constructible Universe model. In order to 'encode truth', one must list out and number all the possible statements of ZFC (since there are countably many, this is possible), and configure a few other technical details.*

Axiom 5 (Rank-into-Rank). *Given a set x, let its* rank *be the largest order type of a sequence of elements $\emptyset \in x_1 \in x_2 \in \ldots \in x$. Let V_λ for an ordinal λ be the collection of all sets whose rank is less than λ. Given models M, N of ZFC, an* elementary embedding *is a function $j : M \to N$ such that (roughly speaking) all statements of ZFC are true in M if and only if they are true in N with the same collection of sets that satisfy the requisite axioms (conceptually, an elementary embedding is a map from one model to another such that everything 'looks' exactly the same from the perspective of all possible ZFC statements). Then there is a nontrivial elementary embedding $j : V_\lambda \to V_\lambda$.*

The last of the above is one of the strongest of what is known as the *large cardinal axioms*, because the ordinal λ must be a cardinal that is very large, whose existence can prove not only the consistency of ZFC but the existence of many other 'large cardinals' as well. (Of course, we don't know if its existence is itself consistent, so that's another question to think about.) Try to think about why a rank-into-rank cardinal λ must also be an inaccessible cardinal.

The Suslin axiom is interesting because one can produce a model of ZFC with a Suslin tree in it by first assuming a much stronger axiom, called the Diamond Principle, which can be used to construct one, and then take a model of that and produce a forcing extension which erases the Diamond Principle but preserves the tree. Thus in order to 'find' the Suslin tree, one looks beyond into a universe in which a certain strong principle holds true, build the tree there, and then force it back into the original universe.

10. Exercises

1. Let the Rational Hypothesis be a hypothesis of the Theory of Steps as follows:

There exists a set Q of numbers such that, $\forall p, q \in Q(\exists r(p < r < q))$.

Prove that the Rational Hypothesis is compatible with the Interstitial Hypothesis by producing a model of Theory of Steps in which both hypotheses are true.