Nonstandard Models

- First-order theories (without finite models) have unintended, non-isomorphic, models. Some are of a different cardinality than the intended model. Others are merely of a different order-type. Either way, the fact that nothing that we can say (in a first-order language) pins down the subject of our mathematical theories (even up to isomorphism) raises questions about the determinacy of basic terms, like ‘uncountable’ and ‘finite’.

Completeness Theorem

- Let us fix on some standard (classical) first-order proof relation, ⊢, and let |= be the standard Tarkian consequence relation (for first-order languages). Then we have:

  - **Soundness Theorem**: If Γ is a set of sentences, and δ is a sentence, then (Γ ⊢ δ) → (Γ |= δ).

  - **Completeness Theorem**: If Γ is a set of sentences, and δ is a sentence, then (Γ |= δ) → (Γ ⊢ δ).

- The proof of Soundness is straightforward. One verifies that each instance of the axiom schemas is valid (i.e., true in all models), and that the inference rule(s) preserve validity.

  - **Note**: Proof of Soundness also amounts to a proof of the consistency of the proof system. If it were inconsistent, then, by Soundness, we would have Ø |= δ and Ø |= ~δ. But, by the definition of |=, ~δ is true in all models if/ if δ is true in none.

- The proof of Completeness is more subtle. The first point is that Completeness means that if Con(Γ & ~δ) then ∃ M |= (Γ & ~δ). (Γ ⊢ δ) means that ~Con(Γ & ~δ) and (Γ |= δ) means that there is no model of (Γ & ~δ). So, if Con(Γ & ~δ), then there is a model of (Γ & ~δ). Thus, to prove Completeness it suffices to prove: Con(Σ) → ∃ M |= δ, for δ ∈ Σ.

  - **Note**: The proof of Completeness will actually show more than this. It will show that if Σ is consistent, then Σ has a countable model. This will generate a puzzle.

- Let us begin with a consistent theory, Σ. Add countably-many constants to the language, c₁, c₂,...,cₙ, and amend the formal system’s properties correspondingly. The constants are called witnesses. Clearly, Σ in the supplemented language, call it Σ⁺, remains consistent.
Note: For simplicity, we assume that the language has only one binary predicate, \( P \). But the technique described generalizes to any predicates of arbitrary arity.

Next, let us enumerate the formulas in the enriched language with some free variable, \( x \), \( \varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots \), and define \( \Phi_n \) to be the formula, \( [(\exists x)\varphi_n(x) \rightarrow \varphi(c_{n^*})] \). The constant \( c_{n^*} \) is the first constant from our enumeration that fails to occur inside any prior \( \varphi \) or \( \Phi \).

We now want to add each \( \Phi_n \) to our theory, \( \Sigma \). The idea is to make sure that whenever our theory proves that there is an \( x \) such that \( \varphi \), it proves this of some \( c \). So, we define:

- \( \Sigma^0 = \Sigma \)
- \( \Sigma^{n+1} = \Sigma^n \cup \{ \Phi_n \} \) (i.e., \( \Sigma^n \cup \{ [(\exists x)\varphi_n(x) \rightarrow \varphi_n(c_{n^*})] \} \))
- \( \Sigma^* = \bigcup \Sigma^n \)

It is clear that each \( \Sigma^n \) is consistent since each new witness acts like a free variable. But any proof is finite, contained in some \( \Sigma^n \). So, if each \( \Sigma^n \) is consistent, then \( \Sigma^* \) must be too.

Lindenbaum’s Lemma: Every consistent theory has a complete and consistent extension in the same language. We proceed as in the propositional case. Let \( \varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \) be an enumeration of all of the sentences in the language of consistent theory, \( \Sigma^\infty \). Then define:

- \( \Sigma_0 = \Sigma^\infty \)
- \( \Sigma_{n+1} = \Sigma_n \cup \{ \varphi_n \} \), if this is consistent, and let \( \Sigma_{n+1} = \Sigma_n \) if not.
- \( \Sigma^* = \bigcup \Sigma_n \) (\( n \in \mathbb{N} \))

Then \( \Sigma^* \) is a complete consistent extension of \( \Sigma^\infty \) (in \( \Sigma^\infty \)'s language) with the features:

- \( \Sigma^* \vdash \delta \) or \( \Sigma^* \vdash \neg \delta \) (because \( \Sigma^* \) is complete)
- \( \Sigma^* \vdash \neg \delta \) if \( f/\Sigma^* \vdash \delta \) (because \( \Sigma^* \) is also consistent)
- \( \Sigma^* \vdash (\delta \& \alpha) \) if \( f/\Sigma^* \vdash \delta \) and \( \Sigma^* \vdash \alpha \)
- \( \Sigma^* \vdash (\exists x)\varphi_n(x) \) if \( f/\Sigma^* \vdash \varphi_n(c_{n^*}) \)

We can now define a model, \( M = \langle U, R \rangle \), of \( \Sigma^* \) by conflating names and their referents.

- \( U = \{ c_1, c_2, \ldots, c_{n^*}, \ldots \} \)

Note: Technically, we identify the elements of \( U \) with equivalence classes. \( c_j \sim c_k \lrarr T \vdash c_j = c_k \). Thus, \( [c_j] \) is the class of \( c_k \)'s, \( [c_j] = \{ c_k : T \vdash c_j = c_k \} \).
- \( R = \{<c_j, c_k>\} \) such that \( \Sigma^* \vdash P(c_j, c_k) \)

- As \( \Sigma \subseteq \Sigma^* \), \( M \models \delta \), for \( \delta \in \Sigma \), this completes the proof.

**Löwenheim–Skolem Theorem**

- So, if \( \Sigma \) is a (syntactically) consistent set of sentences, then it has a **countable** model. This might strike you as perplexing. What if the \( \Sigma \) is (first-order) ZF, which proves Cantor’s Theorem, along with the existence of (a set-theoretic surrogate of) \( \mathbb{N} \) and \( P(\mathbb{N}) \)?

  - **Cantor’s Theorem**: For any set, \( X \), there is no one-to-one correspondence between \( X \) and \( P(X) \).

  - **Proof**: Suppose, for **reductio**, that \( f: X \rightarrow P(X) \) is such a correspondence. Define a subset of \( X \), \( Y \in P(X) \), as follows: \( Y = \{x \in X : x \notin f(x)\} \). Now suppose that there is an \( x \) with \( f(x) = Y \). Then, if \( x \in Y \), \( x \notin f(x) = Y \), which is a contradiction. On the other hand, if \( x \notin Y = f(x) \), then \( x \in Y \), which is also a contradiction. Therefore, the assumption that \( f \) is a one-to-one correspondence between \( X \) and \( P(X) \) is false.

- The standard view is that Cantor’s Theorem shows that cardinality of the powerset of \( X \) is strictly greater than that of \( X \). Hence, the cardinality of the powerset of \( \mathbb{N} \) is **uncountable**.

  - **Skolem**: “By virtue of the [Zermelo] axioms we can prove the existence of higher [uncountable] cardinalities….How can it be, then, that the entire domain…can already be enumerated by means of the finite positive integers [1922, 295]?”

  - **Note**: The standard view is not beyond question. Dummett maintains that “The argument does not show that \([P(\mathbb{N})]\) form[s] a non-denumerable totality unless we assume…that \([P(\mathbb{N})]\) form[s] a **determinate totality** comprising all that we shall ever recognize as a real number: the alternative is to regard the concept of real number as an **indefinitely extensible** one [1993, 27].” This interpretation, however, is objectionably psychologistic. The claim appears to be that we form a ‘definite conception’ of subsets of \( X \) by pairing them with members of \( X \). Then we use that pairing – i.e., that conception – to literally **create** a new subset of \( X \).

- The incongruence of the **Completeness Theorem** and Cantor’s Theorem can be extended.
Löwenheim-Skolem Theorem: If a set of sentences, $\Sigma$ (in a countable language) has a model (where ‘$=$’ means identity), it has a model of any infinite cardinality.

- The ‘downwards’ implication of this can be refined using Mostowski Collapse:
  
  Transitive Submodel Theorem: If $N$ is a transitive model of infinite cardinality $\kappa$, and $\lambda < \kappa$, then there is a transitive submodel of $N$, $M$, such that the cardinality of $\text{dom}(M)$ is $\lambda$, and, for any $\delta$, $N \models \delta$ if and only if $M \models \delta$.

- A model, $M$, is transitive when $x \in \text{dom}(M) \rightarrow x \subseteq \text{dom}(M)$, and a submodel of $N$ when $\text{dom}(M) \subseteq \text{dom}(N)$ and $M$ and $N$ agree on their interpretation of constants, predicates, and function symbols.

  Note: The assumption that there is a transitive model of $\Sigma$ is stronger than the assumption that there is just a model of $\Sigma$!

- Skolem’s ‘paradox’ relies on the assumption of (classical) first-order logic. If we help ourselves to second-order ZF and its standard (non-Henkin) semantics, then the discrepancies above evaporate. But if the theorems are taken to preclude explaining the determinacy of ‘uncountable’, then nothing is gained by resorting to second-order logic. The problem just becomes to explain the determinacy of the second-order quantifiers.

Compactness Theorem

- Theories may have ‘untended’ models of the intended cardinality as well. (Indeed, ZF is not $\kappa$-categorical, it is not the case that all models of ZF of cardinality $\kappa$ are isomorphic.)

  Compactness Theorem: If every finite subset of $\Sigma$ has a model, then $\Sigma$ does too.

  Proof: If $\sim \text{Con}(\Sigma)$, then there is a formula, $\delta$, such that $\Sigma \vdash (\delta \& \sim \delta)$, i.e., some $n \in \mathbb{N}$, and a sequence of formulae, $\theta_1, \theta_2, \ldots, \theta_n$, such that $\theta_n$ is $(\delta \& \sim \delta)$, and each $\theta_i$ is a logical axiom, an element of $\Sigma$, or such that $\theta_i$ follows from previous $\theta$s by a rule of inference. As this list is finite, the number of formulas in it that are also members of $\Sigma$ is. So, $\sim \text{Con}(\Sigma) \rightarrow \sim \text{Con}(\Sigma_{\text{Fin}})$, for some finite subset of $\Sigma$, $\Sigma_{\text{Fin}}$. Equivalently, if $\text{Con}(\Sigma_{\text{Fin}})$, for every finite subset of $\Sigma$, $\Sigma_{\text{Fin}}$, then $\text{Con}(\Sigma)$. Now suppose that every finite subset of $\Sigma$, $\Sigma_{\text{Fin}}$, has a model. Then, by Soundness, every $\Sigma_{\text{Fin}}$ is consistent. As this implies that $\text{Con}(\Sigma)$, $\exists M \models \Sigma$, by Completeness.

Illustration: Consider the following theory in language $\mathcal{L} = \{<\}$. 

• (i) \((\forall x)(\neg (x < x))\)
• (ii) \((\forall x)(\forall y)(\neg ((x < y) \land (y < x)))\)
• (iii) \((\forall x)(\forall y)(\forall z)([(x < y) \land (y < x)] \rightarrow (x < z))\)
• (iv) \((\forall x)(\forall y)(x < y) \lor (y < x) \lor (x = z))\)
• (v) \((\exists x)(\forall y)(\neg (y < x))\)
• (vi) \((\forall x)(\exists y)(x < y) \land (\forall z)(\neg (x < z) \land (z < y))\)
• (vii) \((\forall x)((\exists y)(y < x) \rightarrow (\exists z)(z < x) \land (\forall w)(\neg (w < x) \land (z < w))))\)

• It turns out that sentences, (i) – (vii), completely axiomatize the structure, \(\mathcal{N} = (\mathbb{N}, <)\). For any sentence, \(\delta\), and any model of the sentences, (i) – (vii), \(M, \mathcal{N} |= \delta\) just in case \(M |= \delta\). It follows that the sentences, (i) – (vii), prove \(\delta\) just in case \(\mathcal{N}\) itself is a model of \(\delta\).

  ○ Note: \(\mathcal{N} = (\mathbb{N}, <)\) is not the structure \((\mathbb{N}, <, +, \ast)\), which cannot be axiomatized! But the structure \((\mathbb{N}, <, +)\) can be, as can the real and complex number fields.

• Despite being complete, the sentences (i) – (vii) are not categorical. That is, it is not the case that all models of (i) – (vii) are isomorphic. We can prove this using Compactness.

  ○ Recall: Categoricity is the most that we can hope for. Clearly, if \((\mathbb{N}, <)\) satisfy (i) – (vii), then so does \((\mathbb{N} - \{0\}, <)\), and so on for infinitely-many other structures.

• Argument: Expand the language \(\mathcal{L}\) to include a constant, \(c\). Call the result \(\mathcal{L}^*\). Now consider the sentences (i) – (vii) in tandem with the following infinite set from \(\mathcal{L}^*\).

  ○ \(\Psi_1: (\exists x_1)(x_1 < c)\)
  ○ \(\Psi_2: (\exists x_1)(\exists x_2)((x_1 < x_2) \land (x_2 < c))\)
  ○ ...
  ○ \(\Psi_n: (\exists x_1)(\exists x_2)...(\exists x_n)((x_1 < x_2) \land (x_2 < x_3)...\land (x_n < c))\)
  ○ ...

• Let us take \(\Sigma\) to consist of (i) – (vii) in addition to all \(\Psi_i\) above, and let \(\Sigma' \subseteq \Sigma\) to be any finite subset of \(\Sigma\). Then each \(\Sigma'\) has a model of the form \((\mathbb{N}, <, n \in \mathbb{N})\), where \(n\) is the largest \(n \in \mathbb{N}\) such that \(\Psi_n \in \Sigma'\). (i) – (vii) are all true in \((\mathbb{N}, <, n \in \mathbb{N})\) because they are true in \((\mathbb{N}, <)\), and \(\Psi_n\) is true in \((\mathbb{N}, <, k \in \mathbb{N})\) so long as \(n \leq k\) (where, again, 0 \(\in \mathbb{N}\)). So, by Compactness, (i) – (vii) plus all \(\Psi_i\) must have a model, \(M\), satisfying the same sentence in \(\mathcal{L}\) as \((\mathbb{N}, <)\). But \(M\) is not isomorphic to \((\mathbb{N}, <)\) because the domain of \(M\) contains an object (denoted by \(c\)) with infinitely-many predecessors, while \(\mathbb{N}\) does not.
Note: Although any such model must be complicated (by Tennenbaum's Theorem), a similar argument shows that there is a non-standard countable model of $PA$ (in the language $\{<, +, *\}$). We will see later that such a model could satisfy $PA + \text{‘there is a Gödel number of a proof of a contradiction in } PA\text{’}$. Alternatively, it might be a model of the (non-recursively enumerable) theory, True Arithmetic, i.e., every truth in the language of $\{\mathbb{N}, <, +, *\}$. So, Completeness is not sufficient for categoricity (but categoricity suffices for completeness). Only (first-order) theories with finite models are categorical.

- Details: Any nonstandard model consists of an initial segment that is isomorphic to the standard model, with extra objects ‘tacked on the end’.

- Upshot: Finiteness is not (first-order) definable. If it were, then we could rule out all nonstandard models by conjoining to the axioms of $PA$ the sentence, ‘for all $x$, $x$ has finitely-many predecessors!"

- Clarification: ‘Finite’ is nevertheless absolute for (standard) transitive models, unlike ‘countable’.