Ultrafilters and Non-Standard Models of Arithmetic

Billy Snikkers

September 2022

1 Filters

Definition 1.0.1. A filter on a set X is a subset $\mathcal{F} \subseteq \mathcal{P}(X)$ which decides which subsets should be considered *large*, satisfying

1. $X \in \mathcal{F}$ (X is large)
2. $\emptyset \notin \mathcal{F}$ (the emptyset is not large)
3. If $A \in \mathcal{F}$ and $A \subseteq B$ then $B \in \mathcal{F}$ (supersets of large sets are large)
4. If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$ (large sets have large intersection)

Example 1.0.2.

- The trivial filter: $\mathcal{F} = \{X\}$
- The cofinite filter (X infinite): $\mathcal{F} = \{A \subseteq X \mid X \setminus A \text{ is finite}\}$
- The principal filter generated by $x \in X$: $\mathcal{F} = \{A \subseteq X \mid x \in A\}$

Note that filters have the *finite intersection property* (FIP):

$$\text{If } A_1, \ldots, A_n \in \mathcal{F}, \text{ then } \bigcap_i A_i \neq \emptyset$$

Lemma 1.0.3. *Any subset $S \subseteq \mathcal{P}(X)$ with the FIP has a minimal filter containing it, the filter generated by S*

Proof. Let S' be the closure of S under finite intersections, i.e. the smallest superset of S satisfying 4. Then $\mathcal{F} = \{A \subseteq X \mid \exists B \in S', B \subseteq A\}$ is a filter. S' cannot have introduced the $\emptyset$ by the FIP, and neither could adding supersets, so 2 remains true of $\mathcal{F}$. If $U, V \in \mathcal{F}$, then there are $U', V' \in S'$ with $U' \subseteq U, V' \subseteq V$, so $U \cap V$ is a superset of $U' \cap V'$, which is in S', so $U \cap V \in \mathcal{F}$. 1 and 3 are clear. □

Definition 1.0.4. A filter $\mathcal{F}$ on X is an *ultrafilter* if every subset is either large or *co-large*, i.e.

5. For any $A \subseteq X$, either $A \in F$ or $X \setminus A \in F$.

Principal filters are ultrafilters (since $x \notin A \Leftrightarrow x \in X \setminus A$) but the cofinite filter is not. For example, the subset $\{0, 2, 4, 6 \ldots\}$ of even naturals is neither cofinite or finite, so it is neither large, nor co-large in the cofinite filter on $\mathbb{N}$. Are there any non-principal ultrafilters on an infinite set? Yes, classically. The difficulty in constructing them lies in deciding which of the infinite, but not cofinite, sets are large and which aren't (the cofinite sets are always present in an infinite ultrafilter).

Lemma 1.0.5. *$\mathcal{F}$ is an ultrafilter iff $\bigcup_{i=0}^n A_i \in \mathcal{F}$ implies that $A_i \in \mathcal{F}$ for some i (a large set cannot be a finite union of small sets).*

Proof. Suppose $\bigcup_{i=0}^n A_i \in \mathcal{F}$, but there is no i with $A_i \in \mathcal{F}$. With $B_i = X \setminus A_i$, every such B_i is in $\mathcal{F}$ by the ultrafilter property (5) and therefore $\bigcap_{i=0}^n B_i \in \mathcal{F}$ by repeated application of (4). This set is the complement of $\bigcup_{i=0}^n A_i$, so they cannot both be in the filter without contradicting (5). Hence there must be some $A_i \in \mathcal{F}$. □

Lemma 1.0.6. *(Ultrafilter lemma). Every filter is contained in an ultrafilter.*

Proof. We apply Zorn's lemma to the poset of filters on X containing $\mathcal{F}$. Observe that

- Given a chain of filters $\mathcal{F} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots$, their union is a filter. Axioms (1) - (3) hold for any union of filters, and 4 holds, since for any $A, B \in \bigcup_{i=0}^\infty \mathcal{F}_i$, $A \in \mathcal{F}_j, B \in \mathcal{F}_k$ for some j, k, so if $\mathcal{F}_j \subseteq \mathcal{F}_k$ then A and B both belong to the filter $\mathcal{F}_k$, and so does $A \cap B$.
- A maximal filter is an ultrafilter. If a filter $\overline{\mathcal{F}}$ is not an ultrafilter, then there is some A with $A \notin \overline{\mathcal{F}}$ and $X \setminus A \notin \overline{\mathcal{F}}$. But $\overline{\mathcal{F}} \cup \{A\}$ has the FIP. Any finite intersection involving A can be written $A \cap \bigcap_i U_i$ with each $U_i \in \overline{\mathcal{F}}$. But if this is empty, then $\bigcap_i U_i \subseteq X \setminus A$, contradicting that $\overline{\mathcal{F}}$ is a filter by 3. Hence $\overline{\mathcal{F}}$ is properly contained in the filter generated by $\overline{\mathcal{F}} \cup \{A\}$, so $\overline{\mathcal{F}}$ cannot be maximal.

Zorn's lemma therefore says that there is a maximal such filter among those containing $\mathcal{F}$, since any chain in the poset of such filters is upper bounded by the union filter of the chain. That maximal filter is an ultrafilter containing $\mathcal{F}$. □

Lemma 1.0.7. *Every non-principal ultrafilter contains the cofinite filter.*

Proof. A restatement of this lemma says that non-principal ultrafilters do not have finite sets - the ultrafilter condition forces it to have the complements. Suppose it had a finite set A, and consider the intersection $\bigcap_{a \in A} X \setminus \{a\}$. Every $X \setminus \{a\}$ is in $\mathcal{F}$, since their complements are singletons and $\mathcal{F}$ is non-principal, so the intersection is also $\mathcal{F}$, but that's just $X \setminus A$, contradicting $A \in \mathcal{F}$. □

2 Ultraproducts

Definition 2.0.1. Let $\{M_i \mid i \in I\}$ be a collection of $\mathcal{L}$-structures, and let $\mathcal{U}$ be an ultrafilter on I. The ultraproduct $M = \prod_I M_i/\mathcal{U}$ is an $\mathcal{L}$-structure defined as follows:

- The underlying set is $(\prod_I M_i) / \sim$, where $(a_i) \sim (b_i)$ if they agree on a large set, i.e. $\{i \in I \mid a_i = b_i\} \in \mathcal{U}$.
- If c is a constant symbol, $c^M = [(c^{M_i})]$
- If f is a function symbol of arity n, $f^M([(a_{i1})], \dots, [(a_{in})]) = [(f^{M_i}(a_{i1}, \dots, a_{in}))]$
- If R is a relation symbol of arity n, $([(a_{i1})], \dots, [(a_{in})]) \in R^M$ iff $\{i \in I \mid (a_{i1}, \dots, a_{in}) \in R^{M_i}\} \in \mathcal{U}$.

In other words, the objects of M consist of I-indexed sequences of M_i objects, which are identified if the sequences are *mostly* identical w.r.t. the ultrafilter. Terms are intepreted pointwise, and relation symbols express that two sequences are related if they are related in most of the M_i.

It is an exercise in [1] to prove that the equivalence relation is well-defined, and that interpretation of terms and relations is independent of the choice of representatives.

Theorem 2.0.2. *(Łoś' theorem).*

Let $\{M_i \mid i \in I\}$ be a collection of $\mathcal{L}$-structures, and let $\mathcal{U}$ be an ultrafilter on I. Let $\phi(\overline{x})$ be a first-order formula in the free variables $\overline{x}$, and let $\overline{[(a_i)]}$ be a tuple of elements from the ultraproduct $\prod_I M_i/\mathcal{U}$. Then $\prod_I M_i/\mathcal{U} \vDash \phi(\overline{[(a_i)]})$ if and only if $\{i \in I \mid M_i \vDash \phi(\overline{a_i})\} \in \mathcal{U}$.

In other words, truth in an ultraproduct is characterised by "true in *most* of the M_i", where *most* is understood to w.r.t. the ultrafilter on the indexing set.

Remark 2.0.3. Note that the statement $\prod_I M_i/\mathcal{U} \vDash \phi(\overline{[(a_i)]})$ is non-sensical, as each $[(a_i)]$ (in the tuple $\overline{(a_i)}$) is an object in the domain of the structure $\prod_I M_i/\mathcal{U}$, hence $\phi(\overline{[(a_i)]})$ is not (necessarily) a formula of the language $\mathcal{L}$. There are two options.

1. Let $\mathcal{L}^*$ be the first-order language $\mathcal{L}$ augmented to include a name $c_{(a_i)}$ for every (a_i) in the domain $D_{\prod_i M_i}$ and a name c_{a_j} for every $a_j \in D_{M_i}$, and interpret $c_{(a_i)}$ as the object $[(a_i)]$, and c_{a_j} as the object a_j. Then state Łoś's theorem with each M_i (and the ultraproduct) understood as $\mathcal{L}^*$-structures, with conclusion $\prod_I M_i/\mathcal{U} \vDash \phi(\overline{c_{(a_i)}})$ *if and only if* $\{i \in I \mid M_i \vDash \phi(\overline{c_{a_i}})\} \in \mathcal{U}$

2. Use valuations. *Let* $\{\nu_i : \mathcal{V} \to M_i\}_{i\in I}$ *be an I-indexed family of valuations on the structures* M_i*, and* $\nu : \mathcal{V} \to \prod_I M_i/\mathcal{U}$ *the valuation such that* $\nu(x) = [(\nu_i(x))]$*. For any* $\mathcal{L}$*-formula* φ*,* $\mathcal{M}_\nu(\varphi) = 1$ *if and only if* $\{i \in I \mid M_{i\nu_i}(\varphi) = 1\} \in \mathcal{U}$*.* The idea of the original statement can be recovered by choosing a valuation which picks out the intended objects for each free variable.

The former is clearly the more controversial option, since it involves jamming metatheoretical concepts into the first-order language, stretching the definition of what counts as a "symbol". The author's view is that the spirit of FOL is to cut out a decidable, countable, fragment of objects from the metatheory (ZFC) for the syntax, and then speak of semantic interpretations of that syntax via the metatheory. This is done better by option 2, which leaves it to the metatheory to pick out the tuple of objects $\overline{[(a_i)]}$ and "say ϕ about $\overline{[(a_i)]}$".

Proof. We proceed by induction on φ. If φ is atomic, then it is of the form $R(t_1, \ldots, t_n)$ for some relation symbol R, and terms t_i.

$$\begin{aligned}\mathcal{M}_\nu(R(t_1,\ldots,t_n)) = 1 &\Leftrightarrow \mathcal{M}(R)(\mathcal{M}_\nu(t_1),\ldots,\mathcal{M}_\nu(t_n)) = 1\\ &\Leftrightarrow \{i \in I \mid M_i(R)(M_{i\nu}(t_1),\ldots,M_{i\nu}(t_n)) = 1\} \in \mathcal{U} \quad \text{(by definition)}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(R(t_1,\ldots,t_n)) = 1\} \in \mathcal{U}\end{aligned}$$

If φ is an equality $(t_1 = t_2)$.

$$\begin{aligned}\mathcal{M}_\nu(t_1 = t_2) = 1 &\Leftrightarrow \mathcal{M}_\nu(t_1) = \mathcal{M}_\nu(t_1)\\ &\Leftrightarrow [\mathcal{M}_{i\nu}(t_1)] = [\mathcal{M}_{i\nu}(t_2)]\\ &\Leftrightarrow \{i \in I \mid \mathcal{M}_{i\nu}(t_1) = \mathcal{M}_{i\nu}(t_2)\} \in \mathcal{U} \quad \text{(by definition)}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(t_1 = t_2) = 1\} \in \mathcal{U}\end{aligned}$$

If φ is $\neg\psi$.

$$\begin{aligned}\mathcal{M}_\nu(\neg\psi) = 1 &\Leftrightarrow \mathcal{M}_\nu(\psi) = 0\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi) = 1\} \notin \mathcal{U}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi) = 0\} \in \mathcal{U}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\neg\psi) = 1\} \in \mathcal{U}\end{aligned}$$

If φ is $\psi \wedge \theta$

$$\begin{aligned}\mathcal{M}_\nu(\psi\wedge\theta) = 1 &\Leftrightarrow \mathcal{M}_\nu(\psi) = 1 \text{ and } \mathcal{M}_\nu(\theta) = 1\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi) = 1\} \in \mathcal{U} \text{ and } \{i \in I \mid M_{i\nu}(\theta) = 1\} \in \mathcal{U}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi) = 1\} \cap \{i \in I \mid M_{i\nu}(\theta) = 1\} \in \mathcal{U}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi) = 1 \text{ and } M_{i\nu}(\theta) = 1\} \in \mathcal{U}\\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\psi\wedge\theta) = 1\} \in \mathcal{U}\end{aligned}$$

If φ is $\exists x\psi$

$$\begin{aligned}\mathcal{M}_\nu(\exists x\psi) = 1 &\Leftrightarrow \text{There is some } d \in D_\mathcal{M} \text{ such that } \mathcal{M}_{\nu_{x\mapsto d}}(\psi) = 1 \\ &\Leftrightarrow \text{There is some } d \in D_\mathcal{M} \text{ such that } \{i \in I \mid M_{i\nu_{x\mapsto d_i}}(\psi) = 1\} \in \mathcal{U} \\ &\Leftrightarrow \{i \in I \mid M_{i\nu}(\exists x\psi) = 1\} \in \mathcal{U}\end{aligned}$$

The cases for $\vee$, $\Rightarrow$ and $\forall$ can be obtained by equivalences to formulas using only $\neg, \wedge, \exists$. □

Corollary 2.0.4. *Let M be an $\mathcal{L}$-structure, $\mathcal{U} \subseteq \mathcal{P}(\mathbb{N})$ an ultrafilter on $\mathbb{N}$, and $\mathcal{M} = M^\mathbb{N}/\mathcal{U}$ the ultraproduct of $\mathbb{N}$ copies of M over $\mathcal{U}$ (called an* ultrapower*). For any formula φ and any valuation $\nu : \mathcal{V} \to M^\mathbb{N}/\mathcal{U}$ (which splits into $\{\nu_i : \mathcal{V} \to M_i\}_{i=0}^\infty$)*

$$\mathcal{M}_\nu(\varphi) = 1 \quad \textit{if and only if} \quad \{i \in \mathbb{N} \mid M_{i\nu_i}(\varphi)\} \in \mathcal{U}$$

Example 2.0.5. Let $\mathcal{L}$ be the language of arithmetic and $M = \mathbb{N}$ the "standard-model" $\mathcal{L}$-structure. Consider an ultrapower ${}^*\mathbb{N} = \mathbb{N}^\mathbb{N}/\mathcal{U}$ (for some $\mathcal{U} \subseteq \mathcal{P}(\mathbb{N})$). We call these the *hypernaturals*.

- Each numeral $\overline{n}$ in the language corresponds to $[(n, n, n, n, n, \dots)]$, called a *standard hypernatural*.
- The hypernatural $K = [(0, 1, 2, 3, 4, \dots)]$ is larger than any standard hypernatural, since $\{i \in \mathbb{N} \mid K_i > n\} = \{n + 1, n + 2, \dots\}$, which is co-finite, hence in $\mathcal{U}$.
- The hypernatural $[(p_i)_{i=0}^\infty] = [(2, 3, 5, 7, 11, \dots)]$ (where p_i is the ith prime natural) is prime (any factorisation consists of $[(p_i)]$ and $[(1, 1, 1, \dots)]$), and is also larger than any standard hypernatural.
- The hypernatural $P = [(2, 2 \cdot 3, 2 \cdot 3 \cdot 5, \dots)]$ (where p_i is the ith prime natural) is divisible by every prime standard hypernatural.

Theorem 2.0.6. *For every first-order formula φ in the language of arithmetic, $\mathbb{N} \vDash \varphi$ if and only if ${}^*\mathbb{N} \vDash \varphi$.*

Proof. Suppose $\mathbb{N} \vDash \varphi$, and let $\nu : \mathcal{V} \to {}^*\mathbb{N}$ be a valuation splitting into $\{\nu_i : \mathcal{V} \to \mathbb{N}\}_{i=0}^\infty$. Then $\mathbb{N}_{\nu_i}(\varphi) = 1$ for each of the ν_i (since $\mathbb{N} \vDash \varphi$), so $\{i \in \mathbb{N} \mid \mathbb{N}_{\nu_i}(\varphi) = 1\} = \mathbb{N} \in \mathcal{U}$, giving ${}^*\mathbb{N}_\nu(\varphi) = 1$.

For the converse, suppose ${}^*\mathbb{N} \vDash \varphi$, and let $\nu : \mathcal{V} \to \mathbb{N}$ be a valuation. Construct the valuation ${}^*\nu : \mathcal{V} \to {}^*\mathbb{N}$ which sends $x \mapsto [(\nu(x), \nu(x), \dots)]$. Then ${}^*\mathbb{N}_{{}^*\nu}(\varphi) = 1$, so $\{i \in \mathbb{N} \mid \mathbb{N}_\nu(\varphi) = 1\} \in \mathcal{U}$, and all large sets are non-empty, so picking any i in this set, we have $\mathbb{N}_\nu(\varphi) = 1$. Hence $\mathbb{N} \vDash \varphi$. □

Example 2.0.7.

- There are infinitely many prime hypernaturals larger than $[(p_i)_{i=0}^\infty]$, since $\mathbb{N} \vDash (\forall x)(\exists y)(y > x \wedge \text{isPrime}(y))$, so the same holds for ${}^*\mathbb{N}$.

Corollary 2.0.8. *${}^*\mathbb{N}$ is a model of $\text{Thm}(\mathbb{N}) = \{\varphi \mid \mathbb{N} \vDash \varphi\}$ ("True Arithmetic").*

References

[1] A. Kruckman, *Notes on Ultrafilters.* `https://math.berkeley.edu/~kruckman/ultrafilters.pdf`

[2] "Rising Entropy", *The Ultra Series.* `https://risingentropy.com/the-ultra-series-guide/`