Models of IZF Set Theory

Report on an Alternative Set Theories project

Sam Adam-Day

Introduction

In January 2018, I took part in the Alternative Set Theories project, led by Yurii Khomskii at the ILLC. For this, I investigated models of Intuitionistic Zermelo–Fraenkel set theory (IZF), and looked at their applications to proving separation results for so-called 'omniscience principles'.

IZF is a constructive set theory. It uses intuitionistic logic instead of classical logic; its axioms are mostly those of classical ZF set theory, with a few alterations to avoid validating excluded middle. Its axioms and basic properties are elaborated in [Aczel and Rathjen, 2010, p. 24]. The models of IZF which I investigated are so-called 'full Kripke models'. The description of these models and the proofs of the separation results which make use of them can be found in [Hendtlass and Lubarsky, 2016]. Throughout the following, the metatheory will be classical ZFC.

Full Kripke Models

The goal is to produce a class of models validating all the axioms of IZF which is sufficiently flexible to allow for the creation of models that separate various pairs of omniscience principles (i.e. models in which one of the pair is true while the other false). The usual way of giving a semantics for intuitionistic logical systems is via 'Kripke semantics'. Briefly, a Kripke model for a system in the language $\mathcal{L}$ consists of a partial order $\mathcal{P}$ and for each $\sigma \in \mathcal{P}$ an assignment to a classical structure $(D_\sigma, P_\sigma \mid P \in \mathsf{Predicates}(\mathcal{L}))$, which gives a domain and for each predicate symbol in the language an interpretation of it in that domain, which assignment is moreover monotone: if $\sigma \leqslant \tau$ then $D_\sigma \subseteq D_\tau$ and $P_\sigma \subseteq P_\tau$ for every predicate P (see [Aczel and Rathjen, 2010, p. 20]). This is what happens in full Kripke models, however the construction is somewhat intricate, since one must consider sets whose extensions are expanding, and whose elements are themselves sets whose extensions are expanding, and so on. In order to construct such a model, we begin with a collection of models of classical ZFC, together with a system of elementary embeddings between them.

So, let $\mathcal{P}$ be a partial order. For any $\sigma \in \mathcal{P}$, we let $\mathcal{P}^{\geqslant\sigma}$ be the partial order $\{\tau \in \mathcal{P} \mid \sigma \leqslant \tau\}$. For each $\sigma \in \mathcal{P}$, let M_σ be a model of ZFC (which can be a class in general). Let

$$\left(f_{\sigma\tau} \colon M_\sigma \to M_\tau \mid \sigma, \tau \in \mathcal{P}, \sigma \leqslant \tau\right)$$

be a system of elementary embeddings which coheres in the following sense. For any $\sigma \in \mathcal{P}$, we require that $f_{\sigma\sigma} = \mathsf{id}_{M_\sigma}$, and for any $\sigma \leqslant \tau \leqslant \rho$ that $f_{\tau\rho} \circ f_{\sigma\tau} = f_{\sigma\rho}$. Assume that for every $\sigma \in \mathcal{P}$ we have that $\mathcal{P}^{\geqslant\sigma}$, each M_τ for $\tau \geqslant \sigma$, the assignment $\tau \mapsto M_\tau$ for $\tau \geqslant \sigma$ and $f_{\tau\rho}$ for $\tau, \rho \geqslant \sigma$ are all definable in M_σ; $\mathcal{P}^{\geqslant\sigma}$ should moreover be a set in M_σ.

With these preliminaries in place, we can now proceed to define the full Kripke model, K, over this system. It will consist of a class of objects $K^\sigma \subseteq M_\sigma$ for each $\sigma \in \mathcal{P}$. The objects of K^σ will be particular functions g whose domain is $\mathcal{P}^{\geqslant\sigma}$, and such that for $\tau \in \mathcal{P}^{\geqslant\sigma}$ we have $g(\tau) \subseteq K^\tau$. The idea is that g represents a set whose elements can be different at each node of the partial order; for any $\tau \geqslant \sigma$, the collection $g(\tau)$ is the elements of g at the node τ. For such a function and $\tau \geqslant \sigma$, let $g^{\geqslant\tau} := g|_{\mathcal{P}^{\geqslant\tau}}$; the idea is that $g^{\geqslant\tau}$ is what the set g looks like at the node τ. We will require that the collection of elements of any g only grows as we move up the partial order; this means that if $h \in g(\tau)$ and $\tau \leqslant \rho$ then $h^{\geqslant\rho} \in g(\rho)$.

Now, the collection K^σ is constructed inside M_σ in stages K^σ_α for every α which is an ordinal in M_σ. So work now inside M_σ. Let us assume that we have constructed K^σ_β for every $\beta < \alpha$. For every $\tau \geqslant \sigma$, since $f_{\sigma\tau}\colon M_\sigma \to M_\tau$ is an elementary embedding (which is in M_σ), this means that $K^\tau_\gamma \subseteq M_\tau$ exists for every $\gamma < f_{\sigma\tau}(\alpha)$. Then we let K^σ_α be the set of functions g such that for every $\tau \geqslant \sigma$:

- g has domain $\mathcal{P}^{\geqslant\sigma}$,
- $g^{\geqslant\tau} \in M_\tau$,
- $g(\tau) \subseteq \bigcup_{\gamma < f_{\sigma\tau}(\alpha)} K^\tau_\gamma$,
- if $h \in g(\tau)$ and $\rho > \tau$ then $h^{\geqslant\rho} \in g(\rho)$.

Finally, we let $K^\sigma := \left(\bigcup_{\alpha\in\mathrm{Ord}} K^\sigma_\alpha\right)^{M_\sigma}$.

With the objects of the model thus specified, we may now define truth in the model. This is defined locally for each $\sigma \in \mathcal{P}$, and is given by the following recursion on the complexity of first-order formulae.

$$
\begin{array}{lcl}
\sigma \vDash \bot & \Leftrightarrow & \mathsf{False} \\
\sigma \vDash g \in h & \Leftrightarrow & g^{\geqslant\sigma} \in h(\sigma) \\
\sigma \vDash g = h & \Leftrightarrow & g^{\geqslant\sigma} = h^{\geqslant\sigma} \\
\sigma \vDash \phi \wedge \psi & \Leftrightarrow & \sigma \vDash \phi \text{ and } \sigma \vDash \psi \\
\sigma \vDash \phi \vee \psi & \Leftrightarrow & \sigma \vDash \phi \text{ or } \sigma \vDash \psi \\
\sigma \vDash \phi \to \psi & \Leftrightarrow & \text{for all } \tau \geqslant \sigma \text{ we have } (\tau \vDash \phi \text{ implies } \tau \vDash \psi) \\
\sigma \vDash \forall x \phi & \Leftrightarrow & \text{for all } \tau \geqslant \sigma \text{ and for all } g \in K^\tau \text{ we have } \tau \vDash \phi({}^{g}\!/\!{}_{x}) \\
\sigma \vDash \exists x \phi & \Leftrightarrow & \text{there exists } g \in K^\sigma \text{ such that } \sigma \vDash \phi({}^{g}\!/\!{}_{x})
\end{array}
$$

Then we let $K \vDash \phi$ if $\sigma \vDash \phi$ for all $\sigma \in \mathcal{P}$.

Note the following. (1) The condition $g^{\geqslant\sigma} \in h(\sigma)$ should be understood as follows. h is a set in some $\epsilon \leqslant \sigma$. For every $\delta \geqslant \epsilon$, we have that $h(\delta)$ gives the collection of elements of h at δ; in particular, $h(\sigma)$ is the collections of elements of h at σ. Further, g is a set in some $\lambda \leqslant \sigma$, and $g^{\geqslant\sigma}$ is what g looks like at σ (so $g^{\geqslant\sigma}$ is a function with domain $\mathcal{P}^{\geqslant\sigma}$). Then $g^{\geqslant\sigma} \in h(\sigma)$ expresses that what g looks like at σ is an element of h at σ. (2) Since the interpretation of $\in$ only increases as we move up the model, it is easy to see that the full Kripke model is monotone: if $\sigma \vDash \phi$ and $\tau \geqslant \sigma$ then $\tau \vDash \phi$. (3) The symbols $\wedge, \vee, \exists$ are interpreted as they would be in classical logic. The symbols $\to$ and $\forall$, on the other hand, have a different semantics, requiring the consideration of all nodes above the current. This is where the (Kripke) structure of the partial order becomes relevant, and from where the non-classical properties come. (4) This means in particular that classical logic holds at all terminal nodes (i.e. nodes with no successors). (5) Notice that an interpretation for $\neg$ is not given. Since we are using intuitionistic logic, $\neg$ is not taken as primitive; instead, we define $\neg\phi$ as $\phi \to \bot$. Using the definition above, we can see that $\sigma \vDash \neg\phi$ if and only if for every $\tau \geqslant \sigma$ we have $\tau \nvDash \phi$. (6) The recursive definition does not give us classical equivalences like $(\phi \vee \phi) \leftrightarrow \neg(\neg\phi \wedge \neg\psi)$ and $\exists x\phi \leftrightarrow \neg\forall x\neg\phi$, so we cannot define $\vee$ in terms of $\wedge$ or $\exists$ in terms of $\forall$.

Theorem 1. *The full Kripke model K over any system is a model of IZF; i.e. K satisfies the axioms and rules of inference of intuitionistic logic and for every axiom Φ of IZF, we have $K \vDash \Phi$.*

Sketch Proof. The set existence/construction axioms are all proved in a similar way. We give a set g which at each τ is the object, from the point of view of M_τ, that is to be shown to exist. For example, to show that $\sigma \vDash$ **Emptyset**, we consider the set g with domain $\mathcal{P}^{\geqslant\sigma}$ such that for $\tau \geqslant \sigma$ we have $g(\tau) = \varnothing$. As another example, to show that $\sigma \vDash$ **Infinity**, we let each M_τ construct by induction for each $n \in \mathbb{N}^{M_\tau}$ the set n_τ which is such that $n_\tau(\rho) = \{m_\rho \mid m < n\}$; then the natural numbers set $\mathbb{N}_\sigma$ will be such that $\mathbb{N}_\sigma(\tau) = \{n_\tau \mid n \in \mathbb{N}^{M_\tau}\}$.

Extensionality follows since $\sigma \vDash \forall x(x \in g \leftrightarrow x \in h)$ means, by the interpretation of $\forall$, that from σ upwards g and h are equal as functions, meaning that $g^{\geqslant\sigma} = h^{\geqslant\sigma}$.

Set−Induction is a little tricky. A natural attempt at a proof using the fact that each M_σ is a model of ZFC and therefore of **Foundation** does not work, since we don't assume that every M_σ are well-founded (from the outside). Instead, we must consider a counterexample to **Set−Induction** of minimal rank, and make use of the fact that the system of elementary embeddings above σ is definable in M_σ. If α is the minimal rank of a counterexample in M_σ, then $f_{\sigma\tau}(\alpha) \geqslant \alpha$ is the minimal rank of a counterexample in M_τ; but any counterexample to **Set−Induction** at σ must at some $\tau \geqslant \sigma$ itself contain a counterexample to **Set−Induction**, which will be of lower rank.

Finally to prove **Collection**, we prove that the model satisfies **Reflection**, which implies **Collection**. To do this, we use that each M_σ, being a model of ZFC, satisfies the ZFC Reflection Principle. This together with the correspondence between K^σ_α and V_α in K and in M_σ gives that K satisfies **Reflection**. $\square$

For a full proof see [Hendtlass and Lubarsky, 2016].

Omniscience Principles

With the models in place, we can turn our attention to separation results. Historically, many ways of strengthening intuitionistic logic have been investigated, what are know as *omniscience principles*. The most obvious and 'crude' of these is the famous Law of Excluded Middle.

$$A \vee \neg A \tag{LEM}$$

However there are many other strictly weaker principles, whose addition to intuitionistic logic can be thought of as making it 'less constructive'. Various implications have been shown between this principles. Here we give a few proofs that show the a non-implication relationship.

These 'separation' results are interesting in of themselves, given that they concern fundamental logical principles, but they also have practical applications. When it comes to the project of classifying which theorems of mathematics are non-constructive, one usually seeks to show that the result implies some (known) non-constructive statement. With these separation results, we make available a new proof method for this task: if A and B are omniscience principles such that $A \not\Rightarrow B$, and if we can show that a result P together with A implies B, then P must be non-constructive.

The first weakening of the Law of Excluded Middle comes in the form of the Weak Law of Excluded Middle.

$$\neg\neg A \vee \neg A \tag{WLEM}$$

We form a hierarchy of principles from this by noting that adding it to intuitionistic logic is equivalent to adding De Morgan's Law:

$$\neg(A \wedge B) \to (\neg A \vee \neg B) \tag{DML}$$

We generalise this to produce a series of principles in descending strength:

$$\neg \bigvee_{\substack{i,j<n \\ i\neq j}} A_i \wedge A_j \to \bigvee_{i<n} \neg A_i \tag{WLEM$_n$}$$

Finally, we may continue this process to the ωth level, to obtain:

$$\neg\exists n, m \in \mathbb{N} \colon (n \neq m \wedge A(n) \wedge A(m)) \to \exists n \in \omega \colon \neg A(n) \tag{WLEM$_\omega$}$$

The other two principles which we will be considering involve the determination of truth values on infinite binary sequences. Let $\mathbb{B}$ be the collection of binary sequences on $\mathbb{N}$. The first principle is called the Limited Principle of Omniscience and is due to Brouwer:

$$\forall \alpha \in \mathbb{B} \colon (\forall n \in \mathbb{N} \colon \alpha(n) = 0) \vee (\exists n \in \mathbb{N} \colon \alpha(n) = 1) \tag{LPO}$$

The second principle is a standard weakening of what is known as Markov's Principle: if it is impossible for all terms of $\alpha \in \mathbb{B}$ to be 0, then there exists an n such that $\alpha(n) = 1$. What we will consider here is called Weak Markov's Principle:

$$\forall \alpha \in \mathbb{B} \colon (\forall \beta \in \mathbb{B} \colon (\neg\neg\exists n \in \mathbb{N} \colon \beta(n) = 1 \vee \neg\neg\exists n \in \mathbb{N} \colon (\alpha(n) = 1 \wedge \beta(n) = 0)) \to \exists n \in \mathbb{N} \colon \alpha(n) = 1) \tag{WMP}$$

Separation Results

We are now in a position to prove a couple of separation results.

Theorem 2. *Over IZF, LPO does not imply WLEM_ω.*

We will give two proofs of this result.

First Proof. Let $\mathcal{P}_1$ be the partial order consisting of strings in the alphabet $\mathbb{N}$ of length at most one, ordered by extension (in other words $\mathcal{P}_1$ consists of a bottom node with countably infinitely many successors). Let ϵ be the empty string. Put the universe V at each node and let all the elementary embeddings $V \to V$ be the identity. Let K_1 be the full Kripke model over this system. We will show that $K_1 \vDash \mathrm{LPO}$ but $K_1 \nvDash \mathrm{WLEM}_\omega$.

Assume that $\epsilon \vDash \alpha \in \mathbb{B}$. Then the object $\alpha \in K_1^\epsilon$ can be transformed to a binary sequence $\hat{\alpha}$ (in the metatheory). Since the metatheory is classical, it in particular satisfies LPO, so either $\forall n \in \mathbb{N} \colon \hat{\alpha}(n) = 0$ or $\exists n \in \mathbb{N} \colon \hat{\alpha}(n) = 1$. In the former case we have that $\forall n \in \mathbb{N} \colon \alpha(n) = 0$ holds at every node (since α is a function (in $\mathbb{M}_\epsilon$) and $\mathbb{N}$ is the same in every model, so nothing new can be added to α), so in particular $\epsilon \vDash \forall n \in \mathbb{N} \colon \alpha(n) = 0$. In the latter case $\epsilon \vDash \exists n \in \mathbb{N} \colon \hat{\alpha}(n) = 1$ by definition of truth at ϵ. Hence LPO holds at ϵ. But LPO also holds at all other nodes since they are terminal nodes so obey classical logic. Therefore $K \vDash \mathrm{LPO}$.

To show that $K_1 \nvDash \mathrm{WLEM}_\omega$, we construct infinitely many (distinct) subsets of 1. For $n \in \mathbb{N}$, let $0_n \in K_1^\epsilon$ be such that:

$$0_n(s) = \begin{cases} \varnothing & s = \epsilon, \\ \{\varnothing\} & s = n, \\ \varnothing & \text{otherwise} \end{cases}$$

Note that $\epsilon \vDash 0_n \subseteq 1$.[a] But now the statements "$0 \in 0_n$" are pairwise incompatible, as in the antecedent of WLEM_ω, but there is no $n \in \mathbb{N}$ such that $\epsilon \vDash \neg 0 \in 0_n$ (since for each such n we have $n \vDash 0 \in 0_n$). □

Note that the above proof only gives a 'weak separation': while WLEM_ω doesn't hold in K_1, neither does its negation (since WLEM_ω holds at all the terminal nodes since they obey classical logic). The second proof refines the first, and gives a 'strong separation': a model in which the negation of WLEM_ω holds.

Second Proof. Let $\mathcal{P}_2$ be the partial order of *all* finite length strings in the alphabet $\mathbb{N}$. Put V at each node and let the embeddings be the identity, as before. Let K_2 be the full Kripke model over this system. We have that LPO holds as before, while at each node we can carry out the argument showing that it does not satisfy WLEM_ω. From the latter, we get for each node s that $s \vDash \neg\mathrm{WLEM}_\omega$, which gives the strong separation we desire. □

Theorem 3. *Over IZF, WLEM does not imply WMP.*

Proof. For this we give only a weak separation. Let $\mathcal{P}_3$ be the partial order:

$$\begin{array}{c} \top \\ \uparrow \\ \bot \end{array}$$

Let $f \colon V \to M$ be an elementary embedding of the universe into a model of set theory containing non-standard natural numbers (this can be obtained by taking the ultraproduct of V using a non-principle ultrafilter over ω). Put V at $\bot$ and M at $\top$, and let f be the embedding. Let K_3 be the full Kripke model over this system.

Now, that $K_3 \vDash \mathrm{WLEM}$ is a special case of the fact that any standard Kripke model which is a linear order satisfies WLEM. We will show the general result here. Let K be an standard Kripke model on the linear order $\mathcal{P}$, and take $\sigma \in \mathcal{P}$. Assume that $\sigma \nvDash \neg A$.[b] Then by definition there is $\tau \geqslant \sigma$ such that $\tau \vDash A$. To show that $\sigma \vDash \neg\neg A$, we need to show that for any $\rho \geqslant \sigma$ we have $\rho \nvDash \neg A$. For any such ρ, by linearity

[a] Note that since $0_n \subseteq 1 \equiv \forall x(x \in 0_n \to x \in 1)$, this means that 0_n is a subset of 1 at all nodes in $\mathcal{P}_1^{\geqslant\epsilon}$.

[b] We'll be making use of a classical argument in the metatheory.

there are two cases. (A) If $\rho < \tau$ then we cannot have $\rho \vDash \neg A$ since otherwise by monotonicity $\tau \vDash \neg A$. (B) If $\rho \geqslant \tau$ by monotonicity $\rho \vDash A$ so $\rho \nvDash \neg A$.

To show that $K_3 \nvDash$ WMP, we can take a binary sequence which is 0 on every standard natural number (i.e. every $n \in \mathbb{N}^V$), but is 1 on some non-standard natural number (note that $f[\mathbb{N}^V] \neq \mathbb{N}^M$). The key to seeing why this violates WMP is to notice that its antecedents have double negations, while its consequent does not. In a standard Kripke model $\sigma \vDash \neg\neg\phi$ means that for every $\tau \geqslant \sigma$ there is $\rho \geqslant \tau$ such that $\rho \vDash \phi$. So, given any binary sequence β, either $\top \vDash \exists n \in N\colon \beta(n) = 1$ or it doesn't (using excluded middle in the metatheory), in which case β must be the 0 sequence (since $f[\mathbb{N}^V] \subset \mathbb{N}^M$), which means that $\top \vDash \exists n \in \mathbb{N}\colon (\alpha(n) = 1 \wedge \beta(n) = 0)$. Therefore the antecedent of WMP for α is satisfied at $\perp$. But the consequent is clearly not satisfied at $\perp$. $\square$

Notice that this proof gives only a weak separation: WMP holds at $\top$, so it is not that case that $K_3 \vDash \neg$WMP. The paper [Hendtlass and Lubarsky, 2016] refines this proof and model, giving a strong separation, but the proof makes use of a more elaborate model called the 'Immediate Settling Model'.

Conclusion

The full Kripke model scheme give a versatile tool which allows us to produce quite fine-tined models that satisfy various properties. We have seen here a few examples of the separation results in whose proofs they can be utilised, and have noted the difference between 'strong separation' and 'weak separation' (a uniquely intuitionistic phenomenon). In fact, we have indicated that in both instances in which we found a weak separation, it was possible to refine the model to give a strong separation. It is an interesting research question whether such a process is always possible, and one which I would like to explore further.

References

[Aczel and Rathjen, 2010] Aczel, P. and Rathjen, M. (2010). CST book draft. Available on webpage at `https://www1.maths.leeds.ac.uk/~rathjen/book.pdf`.

[Hendtlass and Lubarsky, 2016] Hendtlass, M. and Lubarsky, R. (2016). Separating fragments of WLEM, LPO, and MP. *Journal of Symbolic Logic*, 81(4):1315–1343.