Inner and Outer Models for Constructive Set Theories

Robert S. Lubarsky
Dept. of Mathematical Sciences
Florida Atlantic University
Boca Raton, FL 33431
Robert.Lubarsky@alum.mit.edu

August 2, 2018

Abstract

We give a set-theoretic presentation of models of constructive set theory. The models are mostly Heyting-valued and Kripke models, and constructions that combine both of those ideas. The focus is on the kinds of constructions that come up in practice when developing models for particular independence results, such as full models, settling, permutation models, and the use of classical generics. We try to convey some of the intuition behind these constructions, such as topological models as forcing a new, generic point, and Kripke models as allowing a change in the underlying universe. The discussion of inner models includes not only permutation sub-models but also L, its basics, and examples of coding constructions from V into L.
keywords: Heyting-valued models, topological models, Kripke models, full models, settling, forcing, L
AMS 2010 MSC: 03C90, 03E35, 03E40, 03E45, 03E70, 03F50, 03H05

1 Introduction

A mathematician who is classically trained, as most are, could well wonder, when confronted with constructivism, what sense it could make. How could Excluded Middle possibly fail, as well as other classical validities? One can give philosophical motivations, say about increasing knowledge over time or the centrality of constructions, or proof-theoretic demonstrations of the underivability of these principles, perhaps by cut-elimination leading to normal form theorems coupled with the observation that some such principle has no normal form proof. Maybe someone could be convinced to accept some anti-classical principle, some principle contradicting classical logic and set theory, such as Church's Thesis or a Brouwerian continuity axiom. Ultimately, those approaches are likely to fail. Most mathematicians will remain unmoved by philosophical arguments, especially regarding their mathematics; proof-theoretic non-derivability is rather formal, and would leave many cold as mere symbol manipulation; anyone wondering how classical logic could fail is unlikely to embrace a principle violating it.

I got a sense of the coherence of anti-classical principles only via model theory. A model falsifying a classical principle was good; one satisfying an anti-classical principle was even better; best of all was a theorem that a model satisfied a principle iff the model was of such-and-such a form. For instance, consider the ultimate principle at play here, Excluded Middle, in its propositional form: $p \vee \neg p$. In the two-node Kripke model, with nodes $\perp$ and $\top$, with p true only at $\top$, $\perp$ will not satisfy $p \vee \neg p$. This seemed like an intuitive example of how classical logic could fail. Its limitation was that classical logic was not false, because the top node $\top$, with no extension, must satisfy classical logic. Considering that $\neg\neg(p \vee \neg p)$ is constructively provable, the best we can hope for by way of satisfying the negation of Excluded Middle would be to model $\neg\forall p(p \vee \neg p)$. That could be done by a kind of iteration of the first model: a Kripke model in which the nodes are indexed by the natural numbers, and at every node there is a proposition which becomes true only at the next node. Indeed, any Kripke model validating $\neg\forall p(p \vee \neg p)$ would have to embed the model described above: there can be no terminal nodes, since they would satisfy classical logic, and every node would have to have an extension at which some proposition p was not true, yet became true at some later extension.

The motivation of this article is to present the kinds of model constructions that are already known, and what they are typically useful for. While the content will always remain about models of constructive theories, in which classical logic fails, it turns out that many can be viewed as variants of well-known classical constructions. In particular, many of the constructions of current-day set theory are forcing extensions of V, and inner models of V, and elementary embeddings which can be viewed as going to simultaneously an extension and an inner model of V. Hence the focus here will be on models in which V embeds, and extensions and inner models thereof.

In practice, there are three known basic techniques for building constructive models: Kripke models, Heyting (or Heyting-valued) models, and realizability. In a realizability model, though, there is no good embedding of V: there is no function $\hat{\cdot}$ such that $\hat{x}$ has the same properties that x does. So here we restrict attention to the former two models and their variants. To be sure, one can easily enough build Kripke and Heyting models which do not embed V, depending upon just what one means by those names; the justification for this choice to study them is that there are standard constructions of Kripke and Heyting models which are usually what one wants, and they naturally embed V. We consider Kripke and Heyting-valued models as two separate techniques, even though the partial order which is an essential part of a Kripke model is a topological space (and hence a Heyting algebra), because in practice there are things you do with Kripke models that you don't do with Heyting models.

It should be observed that the choice of a set-theoretic, ZF-style framework is personal preference, and has little to do with the actual content of the results or constructions. Frequently enough category or type theory is used to build models, and there could be other ways to do it that have yet to be discovered. While there are some differences among what can be expressed most naturally and easily, all these formalisms are powerful and flexible enough that they can simulate each other. One reason to want a set-theoretic development is that ZFC is nominally the gold standard for mathematics, and so something in that style might be more accessible to general mathematicians and in particular to logicians. To anyone well-versed in a different foundational paradigm, it should be easy enough to translate the material here.

The goals here include providing an up-to-date introduction to the subject, to serve as a reference, to introduce some uniform terminology and notation, and to bring together in one place various constructions that are scattered in the literature. One limitation of this work is that it takes insufficient account of some of the earlier literature. In addition to those listed in the references below, researchers such as Beeson, de Jongh, Fitting, H. Friedman, Goodman, Moerdijk, J.R. Moschovakis, Scowcroft, Smorynski, and S. Weinstein, among others, produced work directly relevant to the topic at hand, in some cases current work. A more thorough study would incorporate more of this earlier material.

2 Heyting Models, or Constructive Forcing

From the name Heyting or Heyting-valued model, one might think of any model in which the value of an assertion is a member of a given Heyting algebra. This indeed is the meaning used in [33], where they need such general models to show that Heyting models in their sense form a complete semantics for constructive logic. Historically and practically, though, this is not how Heyting models have been used. Instead, the models that have actually been used have assigned Heyting values to basic properties (like set membership) of only standard objects, as opposed to non-standard objects like infinitesimal reals or non-standard integers, and not imposed any other restrictions on the objects to be considered, resulting in what is called below the *full model.* [1] Once this Heyting-algebra construction was applied to sets, it was quickly realized that this is the constructive analogue of how Boolean algebras are used classically: Boolean-valued models, more commonly called forcing. The essential (and really only) difference between forcing and Heyting models is that in forcing something is true when it is forced densely, whereas in a Heyting model for something to be true it must be forced by the entire space.

By way of an example, consider what must be the simplest forcing partial order of all, $2^{<\omega}$, the set of finite binary sequences. The empty condition forces $\exists n\ G(n) = 1$, where G stands for the generic, because every condition can be extended to a condition with a 1 in it. In fact, the set of such conditions is the entirety of $2^{<\omega}$, save for those nodes of the form 0^n, finite sequences of 0's. If we now view such finite binary sequences as names for open sets within Cantor space, then the open sets forcing a particular occurrence of 1 cover the entire space, except for the one point 0^ω. Because it's not the entire space which is so covered, the empty condition does not force $\exists n\ G(n) = 1$ under the Heyting semantics.

[1] The first topological interpretations of constructive systems [31, 32] were for propositional logic, so the question of the kind of objects allowed was not yet relevant. When extended to predicate logic [27], the objects needed for the examples were standard objects. The first application of topological models to higher-order systems, Scott's interpretation of analysis [29, 30], uses full models up to the type level being considered. Grayson's extension to full set theory [10, 12] is where we first find the full model, as presented here, even though not with that name.

In the remainder of this section, we define full Heyting-valued models, and give examples and intuition for topological models (Heyting models specialized to the opens of a topological space as the Heyting algebra) and non-spatial models (Heyting models that are not topological).

2.1 Full Heyting-valued Models

As described above, the difference between classical and constructive forcing is in the interpretation of truth, and pointedly not in the choice of objects. So the following definition of the members of the full model is identical to that of the terms of the standard forcing language. Furthermore, we can refer to this model as $V[G]$, since it is the smallest model containing the ground model V and the generic G (defined below).

The *(full) Heyting model* over a complete Heyting algebra $\mathcal{H}$ consists of the class of *names* or *terms*, defined inductively by

$$\begin{aligned} V_\alpha[G] &= \bigcup\{\mathcal{P}(\mathcal{H} \times V_\beta[G]) \mid \beta \in \alpha\}, \\ V[G] &= \bigcup_{\alpha \in \mathrm{ORD}} V_\alpha[G]. \end{aligned}$$

Given $\sigma \in V_\alpha[G]$, the meaning of $\langle h, \tau \rangle \in \sigma$ is that the truth-value or degree of truth of $\tau \in \sigma$ is (at least) h. The idea behind calling it the full model is that you throw in absolutely everything you can. Sometimes one does consider inner models of $V[G]$, which we also take to be Heyting models, hence the qualifier of fullness.

The standard embedding $\check{\ }$ of the ground model V into $V[G]$ is defined inductively by

$$\check{a} = \{\langle \top, \check{b} \rangle \mid b \in a\}.$$

A particularly important object, the *generic*, not from the ground model (except in degenerate cases), is given by the name

$$G = \{\langle h, \check{h} \rangle \mid h \in \mathcal{H}\}.$$

The generic is characterised by the equation

$$[\![\check{h} \in G]\!] = h.$$

The semantics of the sentences with parameters is given inductively on the sentences, with the base cases of $\in$ and $=$ given by induction on the ranks of the parameters. There are two equivalent ways of doing this: defining the relation $h \Vdash \phi$, $h \in \mathcal{H}$ and ϕ a sentence, or defining the function $[\![\phi]\!]$ taking values in $\mathcal{H}$. The connection between those two is that $h \Vdash \phi$ iff $h \leq [\![\phi]\!]$. By analogy with the way forcing is usually developed, we do the former.

- $h \Vdash \sigma \in \tau$ iff h is covered in the sense of the Heyting algebra by some $H \subseteq \mathcal{H}$ (that is, $h \leq \bigvee H$), and for all $h' \in H$ there is an $\langle \bar{h}, \bar{\sigma} \rangle \in \tau$ such that $h' \leq \bar{h}$ and $h' \Vdash \sigma = \bar{\sigma}$.

- $h \Vdash \sigma = \tau$ iff for all $\langle \bar{h}, \bar{\sigma} \rangle \in \sigma$, $h \wedge \bar{h} \Vdash \bar{\sigma} \in \tau$, and vice versa.

- $h \Vdash \phi \wedge \psi$ iff $h \Vdash \phi$ and $h \Vdash \psi$.
- $h \Vdash \phi \vee \psi$ iff h is covered by some H, and for each $h' \in H$ either $h' \Vdash \phi$ or $h' \Vdash \psi$.
- $h \Vdash \phi \rightarrow \psi$ iff for all $h' \leq h$ if $h' \Vdash \phi$ then $h' \Vdash \psi$.
- $h \Vdash \perp$ iff $h = \perp$ in $\mathcal{H}$. (Equivalently, $\perp$ can be taken to be $\check{0} = \check{1}$.)
- $h \Vdash \forall x\ \phi(x)$ iff for all σ $h \Vdash \phi(\sigma)$.
- $h \Vdash \exists x\ \phi(x)$ iff h is covered by some H and for all $h' \in H$ there is some σ such that $h' \Vdash \phi(\sigma)$.

As usual, $\neg\phi$ is taken to be $\phi \rightarrow \perp$. To say that a proposition ϕ is *true*, or *satisfied*, in such a model means $1 \Vdash \phi$, otherwise A is said not to be true or to fail. Being false is a stronger property: ϕ is said to be *false* if $\neg\phi$ is true, or equivalently the only value forcing ϕ is $\perp$.

Proposition 1. *The axioms and rules of inference of constructive logic and the equality axioms are all true under this interpretation.*

Theorem 2. *Under IZF, the full model satisfies IZF* [10]*. Under CZF, the full model satisfies CZF* [9]*.*

It is an interesting question which axiom systems are self-realizing; that is, which theories T prove that the full Heyting model satisfies T. [2]

2.2 Topological Models

In most cases, $\mathcal{H}$ is the cHa of the open sets of some topological space $\mathcal{T}$, and G can be considered to be a new element of $\mathcal{T}$. Of course, G is not in the ground model, so it would not be possible for G to be a member of $\check{\mathcal{T}}$. But typically $\mathcal{T}$ has an independent description, and if $\mathcal{T}$ is interpreted in $V[G]$ via that description, then G will be a member of $\mathcal{T}$. I know of no way to make precise the notion of $G \in \mathcal{T}^{V[G]}$ and to prove when that holds, so for now it seems that the only way to convey to the reader what's going on is via examples.

Let's look first at what is no doubt the simplest example, in which $\mathcal{T}$ is taken to be Cantor space, mentioned above. Being quite literal, G is forced to be a set of finite binary sequences. It is easy to see that $\bigcup G$ is forced to be an infinite binary sequence, interdefinable with G (the latter as the set of proper initial segments of $\bigcup G$), so we can identify G and $\bigcup G$. Even though G is not even in the ground model, as an infinite binary sequence it can be taken to be a member of Cantor space as interpreted in the extension.

Another simple example is letting $\mathcal{T}$ be the reals (with the standard topology) [8]. [3] A basic open set is an open interval $\mathcal{I}$, which as a forcing condition means "the generic is in me." Since $\mathbb{R}$ is covered by intervals of arbitrarily

[2] For instance, typically one does not expect that substituting Collection by Replacement will keep a theory self-realizing. The problem is that even if a model satisfies $\forall x \in X\ \exists! y\ \phi(x, y)$, the choice of a term for y may not be unique, so it may not be possible to use Replacement in the meta-theory.

[3] Not only does [8] contain this example, it is the first work in the style of this article, containing many examples of Heyting models and general theorems about them.

small length $\epsilon > 0$, each of which approximates G to within ϵ, the generic is a Dedekind cut. How is this useful? The two standard ways of constructing $\mathbb{R}$ from $\mathbb{Q}$ are Dedekind cuts and Cauchy sequences. It is not hard to see that every Cauchy sequences yields a Dedekind cut. The other direction is a nice exercise for bachelors students. Invariably Excluded Middle is used in that construction, in the form of a case split (because the alternative is to use Countable Choice, and most would automatically avoid that as just too dishonest). The suspicion arises quickly that without EM and CC this converse would not hold. The likely place to look for a counter-example would be a generic Dedekind cut, because as a generic it will have only the properties it is forced to have. As it turns out, this is exactly the case, as the generic over $\mathbb{R}$ is not a Cauchy sequence of rationals.

After this last example, one might ask whether the next step might be to take a generic Cauchy sequence, and whether that gets us anything [19]. What one typically wants from a Cauchy sequence, in order for it to be useful, is a modulus of convergence. That is, for a Cauchy sequence (x_n), for every $\epsilon > 0$ there is a spot beyond which the members of the sequence stay within ϵ of each other. A modulus of convergence is a function that on input $\epsilon = 1/n$ yields such an index. Once one starts to think that the existence of such a modulus might not be provable, the most likely place to look at for a counter-example is a generic Cauchy sequence. A basic open set is a finite sequence of rational numbers, which is an initial segment of the generic, along with an open interval, which constrains the generic in that all further entries in the generic have to come from this interval, as well as the limit of the generic. As expected, the generic has no modulus of convergence. (It bears mention that this is a building block in the construction of a model in which the Cauchy reals, equivalence classes of Cauchy sequences of rationals, are not Cauchy complete. For more on this, see section 6.2 on permutation models.)

Sometimes care must be exercised in understanding G as a member of $\mathcal{T}$, because $\mathcal{T}$ can be understood in classically equal yet constructively different ways. Take for instance forcing with $\mathbb{R}$, as above. That would have to yield a generic real. But what is a real? If the meta-theory is classical, then $\mathbb{R}$ could just as well be taken to be Cauchy sequences as well as Dedekind cuts, so one might be misled to think of the generic as a generic Cauchy sequence. This confusion is not hard to resolve in this case, because what determines the generic is not so much the nature of the points in the topological space being used, but rather the topology on them. For the case at hand, a basic open set $\mathcal{I}$ is naturally partial information about a located cut, with the rationals less than $\mathcal{I}$ being in the lower part and those greater than $\mathcal{I}$ in the upper. In other cases, the right way to view $\mathcal{T}$ is not so clear.

Perhaps the cleanest example of that is letting $\mathcal{T}$ consist of finite subsets of $\mathbb{R}^2$ with the Vietoris topology, which is the topology induced by the Hausdorff metric [20]. To be explicit, a basic open neighborhood of $F \in \mathcal{T}$ consists of an open set containing all of F, as well as finitely many open sets each one of which contains at least one point from F; there is no requirement that these latter open sets must cover F. One might first guess that the generic would be a generic finite subset of $\mathbb{R}^2$. A moment's thought though should make clear that it is not possible to have a new finite subset of a ground model set. Perhaps though the generic is a finite subset of $\mathbb{R}^2$ as interpreted in the generic extension, and since there are new reals then there are also new points in $\mathbb{R}^2$. This also looks problematic, because the next question you should ask is, if the generic is to be a finite subset of $\mathbb{R}^2$, wherever interpreted, what's its size. None stand out as being the natural answer. In fact, when one goes to look for members of G, there aren't any. But G also isn't just the empty set. A clue is afforded by the fact that the topology is induced by the Hausdorff distance function. Each member F of $\mathcal{T}$ determines a distance function $d(z, F)$ for $z = (x, y) \in \mathbb{R}^2$, namely the shortest distance from z to any point in F. Furthermore, F is definable from d. So $\mathcal{T}$ can be viewed as a space of distance functions. As it turns out, that is the right way to view G, as a generic distance function. From this point of view, it's not so surprising that G has no points. That is, using distances given by G, we could triangulate, and determine regions of the plane that look as though they contain members of G. These regions are determined only up to ϵ. So what we cannot do is ever specify how many points are in each such region. There could be any finite number. As such, we can never get our hands on any individual one.

Another such example is the forcing to falsify BD-N [22]. BD-N states that every pseudo-bounded sequence of natural numbers is bounded. The space to do this is the set of bounded sequences, appropriately topologized. What one gets though is not a generic bounded sequence, which is good, because we want something which is not bounded. Rather, the generic is merely pseudo-bounded. This can happen because classically bounded and pseudo-bounded are the same. So it might be unclear at the beginning what the better way is to think about the space, as the bounded or as the pseudo-bounded sequences. The former is easier and more familiar, but as it turns out not the more useful way.

2.3 Non-topological Models

While the opens of a topological space form a Heyting algebra, not every Heyting algebra can be viewed as the opens of a topological space. There are interesting examples of forcing with non-topological or non-spatial Heyting algebras.

One aspect of forcing with spaces is that you can specialize to a point. That is, when examining say a particular object in the generic extension, you can look at what is forced by any neighborhood of a fixed point. As all such neighborhoods overlap, whatever they force must cohere. For example, suppose the term r is forced to be a real number. For any $\epsilon > 0$, the space is covered by opens that determine r to within ϵ. So all of the opens of $\mathcal{T}$ containing a fixed point x are enough to determine r exactly, at x. This is essentially the (well-known) proof that a (Dedekind) real in the extension is given by an arbitrary continuous function from $\mathcal{T}$ to $\mathbb{R}$ in the ground model.

The example of specializing to a point which is of immediate application has to do with the Fan Theorem, which we for now assume holds in the meta-theory. Let T be forced to be a binary tree with no infinite paths. Consider in the meta-theory the tree T' of those binary sequences not forced by any neighborhood of a fixed $x \in \mathcal{T}$ to be out of T. That is, σ is in T' when no neighborhood of x forces σ not to be in T. If T' had an infinite branch B through it, then by the hypothesis on T, some neighborhood of x would force some $\sigma \in B$ to be out of T. But then that violates σ being on T'! So T' has no infinite path. By the Fan Theorem in the meta-theory, T' is finite, which means x has a neighborhood forcing T to be finite. This proves the Fan Theorem in all topological models [8]. Hence, to falsify FT, if a forcing model is possible at all, one needs at least a non-topological Heyting algebra.

As it turns out, this is possible [8]. Let $K(\mathcal{T})$ be the cHa of coperfect open sets of $\mathcal{T}$. This can be viewed either as the subset of the family of open sets consisting of only those open sets with complement a perfect set (with a different join operation), or as the quotient of the opens which identifies $\mathcal{O}$ with $\mathcal{O}\backslash\{x\}$. This is why $K(\mathcal{T})$ is not the opens of a space of points: individual points don't count, in that if a sentence holds everywhere except a point, then it holds everywhere. Letting I be the unit interval, $K(I \times I)$ falsifies the Fan Theorem. (The same paper also examines another non-spatial model, $K(I)$.)

Point-free spaces can be used not only to make the Fan Theorem false, they can be used to make it partially true. The historically first model falsifying the Fan Theorem is based on the Kleene tree, an infinite computable tree with no (infinite) computable branch. Classically, the Kleene tree shows that the computable sets form a model in which Weak König's Lemma is false. Constructively, one can work under recursive realizability, also called Kleene's first model K_1, which (using later developments here) is a model of full IZF set theory based on computability. The (internalization of the) Kleene tree is in this model a non-uniform tree, and the fact that it has no computable paths translates in this model to it having no paths, thereby providing a counter-example to FAN. In fact, this example does even more for us. The computability of the Kleene tree translates into its internalization being decidable (meaning membership in the tree is decidable). So not only is the Fan Theorem violated here, so is a weak fragment of it, the Decidable Fan Theorem, or D-FAN: every decidable bar is uniform. This is interesting, because some fragments of the Fan Theorem, including D-FAN, turn out to be equivalent with other statements that are interesting and natural [14]. Another such fragment is c-FAN, the Fan Theorem for c-bars (where $X \subseteq 2^{<\omega}$ is a c-set if there is a decidable set Y such that $\sigma \in X$ iff every extension of σ from $2^{<\omega}$ is in Y) [3]. Easily, c-FAN implies D-FAN, because every decidable bar is a c-bar. Is the converse implication true? Classical logic is no help here (in our context with sufficient set existence axioms), because that makes all versions of FAN true. Kleene realizability doesn't answer the question, because it makes all versions of FAN false. But it does give us a start. Is there a way to start from K_1 realizability and make D-FAN true which is simultaneously gentle enough to keep c-FAN false?

D-FAN can be stated as "for all decidable $X \subseteq 2^{<\omega}$, if X is a bar then X is uniform." (Easily, if X is uniform then X is a bar.) Plausibly one could make D-FAN true by turning decidable bars uniform – by going to a model with non-standard integers, say, a tree that was infinite could be made uniform by killing it at some non-standard level; this is the approach taken in [25]. A different way to make D-FAN true is to consider decidable trees, which by their decidability can be readily exported into the meta-theory; if they are not uniform internally then they are not uniform externally, which means that they are infinite, which means that they actually do have branches (using WKL in the meta-theory); such a branch could then be included in the model, so that the tree (actually, its complement) no longer represents a bar. The choice of branch to include must be made carefully, to preserve IZF. The easiest way to do that is generically.

One's first guess might well be to force with the tree itself. After all, forcing with a tree like the full binary tree $2^{<\omega}$ produces an infinite path through it. That will not work though in the setting at hand. The trees with which we must be concerned are those like the Kleene tree, in which the terminal nodes are dense (as a moment's thought shows that any tree in which the terminal nodes are not dense already contains a path). Any such partial order forces "$\neg\neg$ the generic path is finite," so this does not help. Furthermore, suppose $\mathcal{T}$ is a topological space forcing "there is an infinite path, say P, through a ground model binary tree T." Then by specializing to a point $x \in \mathcal{T}$ we can build a path through T in the ground model: $\mathcal{T}$ forces that for all n there is a neighborhood of x forcing a necessarily unique sequence of length n to be in P, and stringing these sequences together produces an infinite path. So no spatial forcing will help get paths through Kleene-like trees. The construction that works is to take the tree you want to shoot a path through and to turn it into a Heyting algebra by modding out by all the nodes beyond which the tree is finite [23]. Of course, the forcing just described removes merely one counter-example to D-FAN. To make D-FAN true, all counter-examples must be removed, which can be accomplished by iterating this construction. This is described below in 4.1.

3 Kripke Models

Kripke models can provide a certain flexibility that Heyting models do not, namely the chance to change ground models. That is, in Heyting models, while different statements can become true when going to different Heyting truth values, the objects are always built over some fixed ground model. This ground model might be V, or an inner model like L, or an outer model like a forcing extension $V[G]$, or a constructive model like realizability, but once selected it doesn't change among the various Heyting values. With a Kripke model, the universe in which an object can be said to live can switch from node to node.

Let's illustrate this with an example. Consider the sentence "every Turing machine either converges or diverges." Its failure is most easily modeled with realizability. Is there another way to accomplish that? Heyting models as above will not do: because the integers in such models are standard, con- and divergence there are the same as in the ambient universe. But one can construct a Kripke model counter-example. Let e be (the code for) a machine the convergence of which is not decided by ZFC. (Such a machine exists, lest the Halting Problem be decidable: generate the theorems of ZFC and see what they say about Turing machine convergence.) Easily, $\{e\}$ diverges in V, because convergence would be witnessed by a natural number and hence ZFC-provable. Let M be a model of ZFC in which $\{e\}$ converges and in which V embeds (albeit not elementarily). Such a model can be seen to exist by considering the theory "ZFC + $\{e\}$ converges" along with the diagram of V. (The diagram of V is in the language expanded with a constant symbol for each member of V, and consists of all true atomic and negative atomic assertions.) This theory is consistent by Compactness, and so has a model by Completeness, which we take to be definable over V. Consider the Kripke model with nodes $\perp$ and $\top$, with $\perp < \top$. The structure at $\top$ is M. The structure at $\perp$ is built in V, and will be described in detail shortly as the full model with V at $\perp$ and M at $\top$ [4]; the

[4] This structure does not have an ordinal embedding, as defined below, and so is not covered by the development there. Nonetheless, the definition given there of the full model can be applied in the current case. The only issue is whether this is a model of IZF. It is easy enough to check directly that it is.

intuition is that sets at $\bot$ look like they come from V and can grow arbitrarily as they move to $\top$. At $\bot$, $\{e\}$ does not converge, because the integers are standard, but at $\top$ it does, because there we're in M.

What makes the example above go is that some nodes have only standard sets and others have non-standard sets. This ability to change the ambient universe is the main way Kripke models are flexible and Heyting models not.

A secondary difference between them is that even when the underlying universe is taken to be the same at all nodes of a Kripke model, it is natural to consider sub-models of the full model in terms of the nodes of the Kripke partial order. Typically this happens in the context of some kind of settling, under which sets are not allowed to grow throughout the whole partial order, but must remain constant, or settle down, at some point, examples of which are given below. Furthermore, in the presence of certain kinds of settling, it is natural to consider partial existence: if $p < q$ in the partial order of a Kripke model, there is no expectation that the universe at q is the image of that at p under the transition function; in fact, if the universe at p has all settled down by q, then we will definitely need new sets at q if the model is to violate classical logic. In contrast, the use of settling with topological models is different. It is possible to have immediate settling with partial existence, just as for Kripke models, while using topological ideas (for examples, see sections 4.2.1 and 4.3); but the settling comes into play only because the topology is placed on a tree, where there is a notion of a child, which is an essentially Kripke-style idea. A different kind of topological settling is exposited in 4.2.2; but there the settling can happen without even changing to a different open set, and you can change open sets arbitrarily without settling, an essentially different kind of settling property. Of course, one could consider sub-models of the full topological model in which settling or partial existence is substantive, which is why this distinction is being called secondary. The fact remains that such constructions in the context of topological models have not yet appeared in nature, and so this still seems to be a distinction worth making.

3.1 Full Kripke Models

Analogously to the Heyting models above, there is a notion of a full Kripke model. The presentation below is a refinement of that in [13].

Let $\mathcal{P}$ be a partial order, the elements of which will typically be referred to as nodes. For simplicity, we assume $\mathcal{P}$ has a least element $\bot$, although this is really not necessary. Let $p \mapsto V_p$ be an assignment to nodes of models of ZF. Since these models are typically class models, this assignment cannot be understood as a set of ordered pairs; rather, it is given definitionally. That is, whether x is in V_p, as well as the membership relation and equality, are definable uniformly in x and p.

The full model (over this assignment) will be defined inductively. In order for this induction to work, since we are not assuming that the V_p's are actually well-founded, we need additional structure. An *ordinal embedding* is an order-preserving function f from the ordinals of V_p to those of V_q whenever $p < q$. (So actually f is an indexed set of functions f_{pq}, but the choice of p and q should be clear from the context, so we write it polymorphically as simply f.) Moreover, f must cohere, in that $f_{qr} \circ f_{pq} = f_{pr}$. Finally, whenever $p < r$, there is a finite sequence $p = q_0 < q_1 < ... < q_n = r$ such that each $f_{q_i q_{i+1}}$ is either an isomorphism or an elementary embedding. We say that such an assignment is *admissible* if for all p the entire structure beyond p (namely the set $\mathcal{P}^{\geq p}$, the assignment $q \mapsto V_q$ $(q \geq p)$, and the restriction of f to $\mathcal{P}^{\geq p}$ (i.e. the family f_{qr}, $r > q \geq p$)) is definable in V_p. Without loss of generality we will always assume that $V_\bot = V$, since we could confine ourselves to working in $V_\bot$ anyway.

Example 3. *Let $\mathcal{P}$ be finite, and if $p < q$ choose V_q to be a specific inner model of V_p. For instance, $\mathcal{P} = \{\bot, \top\}, V_\bot = V[G]$ for some forcing generic G, and $V_\top = V$. In this case, the ordinal embedding can be taken to be the identity function. Or let $V_\bot = V$ and let $V_\top$ be some ultrapower of V by an ultrafilter in V. Here the ordinal embedding would be the elementary ultrapower embedding. (In practice, the ultrafilters will not be countably complete, so that the ultrapower has non-standard integers.) For an example with $\mathcal{P}$ infinite, let $\mathcal{P}$ be ω and each V_n be V, with f as the identity.*

Example 4. *For an example of how admissibility could fail, let $V_\bot = V$ and $V_\top = V[G]$, where $\bot < \top$ ($V[G]$ is not definable in V). Or let $\mathcal{P}$ be ω, V_0 be V, V_1 be the ultrapower of V by some non-countably complete ultrafilter $\mathcal{U}$, and more generally V_{n+1} be the ultrapower of V_n by the image of $\mathcal{U}$ in V_n. The problem here is that $\mathcal{P}$ is not a set in any V_n once $n > 0$.*

Remark 5. *Even though the first counter-example of V going to $V[G]$ is being excluded here, it does speak to a major intuition behind constructivism, namely that of knowing more as time goes on. There are still ways to build a model in that situation though, just of a different flavor from the full model as defined below, so they will not be considered here. Also, one might well ask why the various V_p's are to be models of ZF and not of IZF. The reason IZF models are disallowed here is that they would bring up all sorts of additional issues. For one, there are various kinds of models of constructive theories, whereas classical theories have only one notion of a model. Beyond that, all of the constructive models have additional structure – realizability has realizers, topological models have open sets, Kripke models have nodes – which would have to be accounted for in the construction to come, adding to the complication. Again, it is possible to deal with this situation. For an example that simultaneously deals with both of these issues – later nodes being IZF models properly extending earlier nodes – see* [23].

Given an admissible assignment over the partial order $\mathcal{P}$, we define the full model M over it. Note that M depends on $\mathcal{P}$, the assignment V_p, and the embeddings f, mention of which is suppressed in the simple notation M. At node p, the universe M^p will be the union of the M^p_α's as α ranges over the ordinals of V_p. In addition, the transition function k^{pq} from M^p to M^q $(q > p)$ will be defined as the union of the partial transition functions k^{pq}_α defined along the way, from M^p_α to $M^q_{f(\alpha)}$. Since these partial transition functions cohere, we will drop the mention of α. Similarly, we do not mention p and q and refer just to k, allowing k to act polymorphically.

Definition 1. *M^p_α consists of the functions g with the following properties:*

- *$dom(g) = \mathcal{P}^{\geq p}$,*
- *$g \restriction \mathcal{P}^{\geq q} \in V_q$,*

- *$g(q) \subseteq \bigcup_{\beta<f(\alpha)} M^q_\beta$, and*
- *if $h \in g(q)$ and $q < r$ then $k(h) \in g(r)$.*

The transition function k works by restricting the domain.

The definition just given is an induction on ordinals in various models, which could be non-standard, and so needs further explanation beyond what was offered in [13]. There is no problem working at node p inductively on the ordinals of V_p, even if those are ill-founded, because the induction is taking place within V_p. However, the inductive definition given refers also to M^q_β, for q possibly strictly extending p and $\beta < f(\alpha)$ in V_q. If the ordinals are the same in V_p and V_q this won't be a problem, as the induction in V_p applies just as well to V_q. Consider though the case in which f_{pq} is an elementary embedding. Within V_p, an induction at stage α could well assume that M^p_β is well-defined for all $\beta < \alpha$, but it is at best unclear how it would be legitimate to assume well-definedness of M^q_β for all $\beta < f(\alpha)$ in V_q. It does not do to use elementarity on the assertion "for all $\beta < \alpha, M^p_\beta$ is well-defined." Yes, superficially it looks like applying f yields "for all $\beta < f(\alpha), M^q_\beta$ is well-defined," which is what we want. But recall what the notation M suppresses: M^p_β is an abbreviation for "the full Kripke model, up to height β, based on $\mathcal{P}^{\geq p}$ and the assignment V_q to $q \in \mathcal{P}^{\geq p}$ and f." With that realization, applying f to "for all $\beta < \alpha, M^p_\beta$ is well-defined" would yield the well-definedness up to $f(\alpha)$ of the full model based on f applied to the system $\mathcal{P}^{\geq p}$ and its assignments (and even that much only if we extended the notion of f from an elementary embedding on the ordinals to one on all of V_p, or at least on enough of V_p to include $\mathcal{P}^{\geq p}$). There is no reason to think that $f(\mathcal{P}^{\geq p})$ is $\mathcal{P}^{\geq q}$ and other such problems. Rather, one must quantify over all partial orders and embeddings too. This same problem, incidentally, also comes up in the proof of $\in$-induction.

Lemma 6. *For all ordinals $\alpha \in V$ and partial orders $\mathcal{P}$, with least element $\perp$ and $ORD^{V_\perp} = ORD^V$, and all admissible ordinal embeddings f, $M^\perp_\alpha$ is defined.*

Proof. Assume inductively that for all $\beta < \alpha, \mathcal{P}$, and f, $M^\perp_\beta$ is defined. Let $\mathcal{Q}$ be a partial order and g an admissible ordinal embedding on the system $V_q, q \in \mathcal{Q}$. For $q \in \mathcal{Q}$, let $\perp = p_0 < p_1 < ... < p_n = q$ be as given by the admissible assignment. We show inductively on $i \leq n$ that within V_{p_i}, for all $\beta < g(\alpha), \mathcal{P}$, and f, $M^\perp_\beta$ is defined. For $i = 0$, $g(\alpha) = \alpha$, and this is just the inductive hypothesis. Given the inductive assertion for a value $i < n$, $g_{p_i p_{i+1}}$ is either an isomorphism or an elementary embedding. In the former case, if within $V_{p_{i+1}}$ there were some counter-example $\beta < g(\alpha), \mathcal{P}$, and f, that would also be a counter-example within V_{p_i}, because $V_{p_{i+1}}$ is definable within V_{p_i}. In the latter case, the elementarity of g transfers the truth of the statement about $g(\alpha)$ from V_{p_i} to $V_{p_{i+1}}$. So for each $q \in \mathcal{Q}$, within V_q, for all $\beta < g(\alpha), \mathcal{P}$, and f, $M^\perp_\beta$ is defined. In particular, we can take $\mathcal{P}$ to be $\mathcal{Q}^{\geq q}$ and f to be g. For those choices, the interpretation within V_q of $M^\perp_\beta$ is just M^q_β, where M is based on $\mathcal{Q}$ and g. This is all that is needed for $M^\perp_\alpha$ to be defined.

□

Theorem 7. *The full model satisfies IZF.*

Proof. The proof is basically the same as in [13], even if the context there is more limited. Namely, in [13], the embeddings f_{pq} all had to be elementary embeddings from the entire model, as opposed to here, where they are allowed to be elementary embeddings of the ordinals only, or the identity function. That makes no difference in the proof, except in the case of $\in$-Induction, which we now show.

Suppose that $p \models \forall x\ (\forall y \in x\ \phi(y) \to \phi(x))$; we need to show that $p \models \forall x\ \phi(x)$. If not, then for some $q \geq p$ and $g \in M^q$, $q \not\models \phi(g)$. This state of affairs is a true statement in V_q: V_q satisfies "there is a partial order $\mathcal{Q}$, with bottom element $\perp$ and admissible ordinal assignment f and full model K, satisfying $\forall x\ (\forall y \in x\ \phi(y) \to \phi(x))$;, with a counter-example g in $K^\perp$." Because V_q is a model of ZFC, and $K^\perp$ is defined via an induction along the ordinals of V_q, there is an example $\mathcal{Q}, f, K$, and g, with g of least possible rank, say α, among all such models. In this model, letting x from above be g, $\perp \models \forall y \in g\ \phi(y) \to \phi(g)$. Since g is a counter-example, there is an $r \geq \perp$ and an $h \in K^r$ such that $r \models h \in g$ yet $r \not\models \phi(h)$. Let $\perp = q_0 < q_1 < ... < q_n = r$ be as given by the admissible assignment. We show inductively on $i \leq n$ that $f(\alpha)$ is the least rank of a counter-example within V_{q_i} to induction for ϕ. For $i = 0$, $f(\alpha) = \alpha$ was chosen as the least such rank. For the inductive step, given the inductive hypothesis for a value $i < n$, $f_{q_i q_{i+1}}$ is either an isomorphism or an elementary embedding. In the former case, if $f(\alpha) \in V_{q_{i+1}}$ were not the least rank of a counter-example, then restricting everything ($\mathcal{Q}$ and the assignment) to the nodes q_{i+1} and above, we would have a counter-example of rank smaller than $f(\alpha)$ within V_{q_i}, contradicting the inductive hypothesis. In the latter case, the elementarity of f transfers the minimality of $f(\alpha)$ from V_{q_i} to $V_{q_{i+1}}$. Finally, letting i be n, $f(\alpha)$ is the least ordinal rank in V_r of a counter-example to ϕ. But g has rank $f(\alpha)$, and so h as a member of g has a smaller rank. This is a contradiction. $\square$

In the following examples, M is an ultrapower of V via a non-countably complete ultrafilter. That means V embeds elementarily into M, and M has non-standard integers.

Example 8. *WLPO does not imply MP* [13]*: The Limited Principle of Omniscience states that every (binary) sequence either is all 0's or has a 1. Markov's Principle states that if a sequence is not all 0's then it has a 1. Trivially, LPO implies MP. Weak LPO is a slight weakening of LPO, that every sequence either is all 0's or it's not. To see that WLPO does not imply MP, let $\mathcal{P}$ consist of $\perp < \top$, and assign V to $\perp$ and M to $\top$. The full model over that structure satisfies WLPO but not MP. (Consider a sequence which is 0 at all standard places and has a 1 at a non-standard place.)*

Example 9. *WKL does not imply WLPO* [13]*: Weak König's Lemma states that every infinite (binary) tree has an infinite branch. It is not hard to see that WLPO implies WKL: using WLPO, any tree restricted to the descendants of a node is either finite or infinite, which can be used straightforwardly to build a branch. To see that the reverses implication fails, let $\mathcal{P}$ consist of a bottom node $\perp$ and two incomparable successors $\top_0$ and $\top_1$. Assign V to $\perp$ and to $\top_0$ and M to $\top_1$. The full model satisfies WKL: if $\perp \Vdash T$ is an infinite tree, then at $\top_1$ T is interpreted as an infinite tree T_M in M; letting B be an infinite branch through T_M, the object in the Kripke model which looks at $\perp$ and $\top_0$*

like B restricted to the standard natural numbers and at $\top_1$ like just B is the infinite branch desired. WLPO fails, by considering a sequence which is all 0's at standard places and has a 1 in a non-standard place.

3.2 Settling

What characterizes the full model is that sets can keep growing throughout the partial order. Under settling, the sets have to stop growing at some point. Within this intuition, a distinction can be made with respect to what kind of object the set has to settle down to, an external set (say something in V, or more precisely the internalization of such) or an internal set (something in the model currently being constructed which may not come from V). Another way to view this distinction is whether, once a set has settled down, its members have also settled down (yes in the former case, no or, more accurately, not necessarily in the latter).

3.2.1 Settling to External Sets

We describe this kind of settling via two examples.

Example 10. *Class-based settling: Let $\mathcal{P}$ be the class of ordinals. Of course, one cannot think of an object in the Kripke model as a function in the standard sense with domain $\mathcal{P}$, because such a creature would have to be a proper class. But one could consider a function as being given by a definition. For instance, the function which at node κ looks like κ is an ordinal which is not the image $\check{\alpha}$ of any ordinal α from V. In this sense, we could speak of the full model over ORD with V assigned to each node. The model of interest now though is not this full model, but rather the one consisting of actual set-sized functions, with domain some ordinal κ. Beyond κ the set represented by this function does not change; another way to look at it is that it is the image $\check{x}$ of a ground model set $x \in V$. The Kripke set has settled down by κ. The reason to consider this model is that it shows that CZF does not prove Power Set. In fact, full Separation holds in this model, so it shows that IZF - Power Set + Subset Collection does not prove Power Set.* [18]

Example 11. *Class-based settling to an inner model: This example is a lot like the previous one, with $\mathcal{P}$ being ORD and sets settling down by some ordinal to something in V, only here the functions are from $V[G]$, where G is generic for Cohen forcing over V. Normally Cohen forcing is thought of as giving a subset of $\mathbb{N}$, but by identifying $\mathbb{N}$ with $\mathbb{N} \times \mathbb{N}$, G can be thought of as a relation on $\mathbb{N}$. This model shows that IZF - Power Set + Exponentiation does not prove Subset Collection.* [18]

It bears observation that this kind of settling produces models which violate Power Set, by design.

3.2.2 Settling to Internal Sets, and Immediate Settling

The external settling models above violate Power Set, and indeed such models must do so (in all non-trivial cases). Consider a set X as a possible power set of 1. If X has settled down to an external set by node p, then the only subsets of 1 that X could contain there are 0 and 1. If p has some extension then any non-trivial model will have a set which looks like 0 at p but then ends up being 1 at some extension, witnessing that X is not the power set of 1 at p. So to model Power Set, a different kind of settling is needed.

We would like to introduce this settling and its application via an example. Either of the examples from the section on full Kripke models would do; for specificity, we will consider the first, the two-node model separating WLPO and MP. In this model, MP fails, but MP is not false. At $\top$, classical logic holds. This is called a *weak separation*. To get a *strong separation*, one in which MP is false, we would like to iterate the construction. Your first guess as to how to do this might well be to assign V to $\bot$, M to $\top$, and however you went from V to M (say via an ultrafilter $\mathcal{U}$), do that to M (say via $f(\mathcal{U})$) to get a model non-standard relative to M, place that model at some successor of $\top$, and then iterate this procedure through ω. With this assignment, you would want to take the full model. Indeed, WLPO would be true there, just because the partial order is linear, and MP would fail, because you're always getting new non-standard integers. The problem is in defining the model. Admissibility is lost. If the partial order is the standard ω, then it does not exist in any of the associated models after $\bot$'s V. One could try to piece together the non-standard ω's that appear along the way, but this is starting to get complicated.

A simpler approach is just to use immediate settling to an internal set [13]: instead of taking the full model over ω, allow only those sets that settle down at the node after they appear. At that next node, there are new sets; that is, sets that are not in the range of the transition function from previous nodes. Those new sets can then grow at the node after that, but then they would have to settle down. This solves the problem of how the partial order, in this case ω, can get away with not being in the base models: all the information needed to be built into a set is how it changes once.

Of course, one is then left with the question of how Power Set could hold. After all, how could X ever be the power set of, say, $1 = \{0\}$, if X must eventually settle down, yet new subsets of 1 keep on being introduced? The answer is that X settles down to an internal set. Take the example above, where $\mathcal{P}$ is ω. At any node $n \in \omega$, what is the power set of $1 = \{0\}$? Viewed externally, at n, 1 has three subsets, namely $0 = \emptyset$, 1 itself, and the set that looks like 0 at n and then grows to 1 at $n+1$, which we call $1_\top$. So at n, this power set looks like $\{0, 1, 1_\top\}$. Under the transition function, 0 goes to 0, 1 to 1, and $1_\top$ also goes to 1. But then a new $1_\top$ appears, and the power set remains settled, still having three elements, one of which in some sense was already in the power set at n and in another sense is new. The set $\{0, 1, 1_\top\}$ is not the internalization of an external set, but it is still settled if it goes to itself at the next node, even though not all of its members are settled yet.

With regard to the technical details, just as with the full Kripke model above, the exposition in [13] needs some refinement to be correct. The problem, as before, is getting the induction right in a context where there may be non-standard models. The additional challenge in thinking of this as a Kripke model is that the underlying partial order, say ω as in the example above, is typically not in any of the V_p's (except $V_\bot = V$) since they are ω-non-standard. The solution makes essential use of the fact that the settling is immediate, so that the model and the semantics can be defined locally, with reference to only the immediate successor nodes. This will appear in [1].

3.2.3 Uniform and Non-uniform Settling

Immediate settling is just a special case of the more general uniform settling. Under uniform settling, one starts with a partial order $\mathcal{P}$ in which the terminal nodes are dense (every node has an extension which is terminal). Then one places a copy of $\mathcal{P}$ (actually, $f(\mathcal{P})$, the image of $\mathcal{P}$ under the elementary embedding) at each such terminal node, and iterates that procedure ω-many times. Immediate settling is what you get when $\mathcal{P}$ consists merely of $\bot$ followed by a set of children (a tree of height 1).

For non-uniform settling, consider the examples given for external settling (Sec. 3.2.1). Every set there settles down for sure, but the settling is not uniform: the objects at node 0 settle down at all possible ordinals. In contrast, the discussion of internal settling was about only uniform settling: every node p has a set Q of successors such that all sets at p must settle down by all $q \in Q$. One might thereby make the mistake of identifying external with non-uniform settling and internal with uniform. We see no reason for this identification to be valid. Uniformity seems to be orthogonal to internality, in that there could be internal non-uniform models and external uniform ones. Consider for instance the Kripke model based on the partial order ω, with V associated to each node. Take all those sets that settle down to internal sets anywhere along the way. This is an example of non-uniform settling to internal sets. Also, one can instead take those sets that settle down the node after they are introduced to something in V, for uniform settling to external sets. It may not be clear what holds in these models or why someone would be interested in them; the point remains, they are perfectly legitimate models. We leave development of this subject to future work.

3.3 Sideways Settling

There is a different kind of settling that can be useful ([13], Theorem 5.7). Consider the partial order with bottom node $\bot$ and children n for $n \in \omega$. Associate V with $\bot$, and the same M with each n. Take the submodel of the full model of those g's that are eventually constant (i.e. for some n and all $i > n$, $g(i) = g(n)$).

4 Heyting-Kripke Models

There are constructions that use a mix of ideas from both Heyting and Kripke models. It would be interesting to know whether it is necessary to have such a mix to answer the questions for which these mixed models were developed. It would also be interesting to have some general framework into which these constructions could be placed. Here we content ourselves with just describing some examples.

4.1 Iterating Heyting Models

Heyting models are the constructive version of forcing. An important forcing technique is iteration. What is the constructive analogue of iteration?

Iteration is used when you have to do more than one forcing. If all the partial orders of concern are in the ground model, then a simpler form, product forcing, suffices. If instead some of the p.o.s needed for the forcing are only in the generic model for earlier forcings, then an actual iteration is necessary. Doing either a product or an iteration finitely often is unproblematic, constructively as well as classically. Issues arise only with an infinite number of forcings, which are naturally arranged along the order-type of some ordinal. Mostly it's a question of what to do at limits: at what places along this ordinal should a condition be allowed to be non-trivial? All of them? Only finitely many? Or what? Classically, the decision often involves set-theoretic concepts, such as countability, inaccessibility, or stationarity. Needless to say, these solutions are problematic constructively, even starting from the idea of organizing the forcings along a linear ordinal, to say nothing of the other, more advanced concepts.

Fortunately, the use of limits can typically be finessed constructively. Iteration comes up when you want a model satisfying an assertion of the form "every structure with property A has property B." This would typically come up when B implies A, and A and B are close, so they might be equivalent. Suppose you had a forcing for any structure not satisfying B that would make it not satisfy A. Then you could just do those forcings, one after the other, for each such structure, including those that come up along the way, until the process closes off. At that point, you'd be left with a model in which the only structures left satisfying A are those you couldn't force not to have A, namely those with B. If your context is constructive, though, you don't actually have to do the forcing to kill A. It's enough to threaten to do so. If you're in a Kripke model, and some later node does the forcing to kill A, then at the current node it is false that A holds, since A fails later; this suffices to have the assertion not apply to the structure at hand.

A concrete example might be useful. Recall that D-FAN is "for all decidable $X \subseteq 2^{<\omega}$, if X is a bar then X is uniform." In 2.3 we described how to force to get a decidable, non-uniform set of nodes to be not a bar. To get D-FAN false, this forcing would in some sense have to be applied to all decidable, non-uniform sets, which seems to call for some kind of iteration. In [23] this iteration was organized as a Kripke-like model. To the bottom node is associated the K_1-realizability model. At any node p, the children are indexed by those Heyting algebras forced at p to have been constructed as above from a decidable, non-uniform tree, along with a value from this Heyting algebra. Over this partial order with associated models, take the full model. As above, D-FAN holds, because if at a node p a decidable tree is not uniform, some later node forces a path through it, so at p the tree couldn't have represented a bar.

What keeps this from being just a Kripke model is the semantics. For instance, in a Kripke model, a disjunction is true at a node exactly one when of the disjuncts is true at that node. In the current model, topological (or Heyting) concepts play a role, in that there is a notion of a set of nodes covering a node, just as how the join of a set of Heyting values can be greater than another Heyting value. So in this model, a node validates a disjunction if it has a cover each member of which validates one of the disjuncts. Similar considerations apply to membership and the existential quantifier. It should not be a surprise that some such consideration is necessary, because the nodes came from Heyting algebras, yet if there were no consideration of such covering the Heyting structure would be completely lost.

It should be mentioned even if only briefly that due respect much be shown to the computational (K_1) nature of the setting. This does influence the construction of the model: the covers used in the semantics are stricter than arbitrary covers for the Heyting algebras, to make them computational.

4.2 Topological Settling

Settling is an idea that more naturally fits with Kripke models. The basic intuition behind settling is that there is some future stage at which some qualitative change happens. This is consistent in style with a Kripke model on a discrete order, such as a tree, where there is a notion of a next step, or in the case of non-uniform settling a notion of a step; it is inconsistent in style with the sense behind a topological space, where usually an increase in information is thought of as an open set gradually or continuously getting smaller. Of course one can come up with counter-examples to formal versions of those assertions, and we will even see some below. We still think this is fair as a stylistic description, one that is usually true. Coupled with the fact that the examples to be given were developed using a mix of topological and Kripke intuitions, it seems right to include this topic in the category of Heyting-Kripke models.

4.2.1 Topological Settling to Internal Sets

Suppose you had a topological space in which some singleton sets $\{x\}$ were open. In fact, suppose that the set of such points was dense. Then any classically valid assertion you might want to falsify, even if not true in the induced topological model, would not be false: it would be not not true. This would be a prime candidate for iteration, which in this case would be placing another copy of the space at each of these singleton opens, or some variant thereof. This is starting to look like the uniform settling discussed briefly above. Where such a construction has actually come up already, models 11 and 18 and theorem 5.7 of [13], the space was simple, and so ended up looking more like immediate settling. We describe first an oversimplified version of these models, one that did not even need to appear in [13], for expository purposes, and then describe the modifications needed for the other models.

Let $\mathcal{U}$ be a non-principal ultrafilter on ω. Let $\mathcal{T}$ be the space with points $\omega \cup \{*\}$, the discrete topology on ω, and open neighborhoods of $*$ sets of the form $\{*\} \cup X$, where $X \in \mathcal{U}$. Perhaps it is useful to try to view the full topological over $\mathcal{T}$ as a Kripke-like model. Let $\mathcal{P}$ have bottom node $*$ and successor nodes $n, n \in \omega$. Then the universe at n is simply V, and at $*$, inductively, the universe consists of functions g with domain $\mathcal{T}$ such that if $h \in g(*)$ then h is also in the universe at $*$, and $\{n \mid h(n) \in g(n)\} \in \mathcal{U}$. So this is like a Kripke model, only that, even if something is a member of g at the bottom node $*$, that fact can be forgotten at some later node; all that must happen is for that membership fact to hold at ultrafilter-many later nodes. (Of course, this member might itself be interpreted differently at various later nodes.)

It is a simple matter to iterate this topology: place a copy of $\mathcal{T}$ at each open point n. This can be formalized as follows. Consider the tree $\omega^{<\omega}$, finite sequences of natural numbers. A basic open set $\mathcal{O}$ contains a unique shortest sequence σ, and for all $\tau \in \mathcal{O}, \{n \mid \tau^\frown n \in \mathcal{O}\} \in \mathcal{U}$. The idea behind the immediate settling model is that at every node you place a copy of the full model on $\mathcal{T}$. So at a node σ, you have sets that can change when going to a child $\sigma^\frown n$ (as long as they respect the topology when doing so), at which point they have all settled down; but then there are new sets that themselves can change at the next nodes.

We now give the formal details. Since every node looks exactly like every other node, we can dispense with the tree of sequences, and work just with the space $\mathcal{T}$. The sets g are defined inductively, as are the functions k_i from the model to itself, which play the role of transition functions in a Kripke model from $*$ to i. Then g is a set in this model if

- g is a function with domain $\mathcal{T}$,
- $g(*)$ is a collection of sets from this model, as is $g(i)$ (where $i \in \omega$),
- if $h \in g(i)$ then $k_j(h) \in g(i)$ $(i, j \in \omega)$, and
- if $h \in g(*)$ then $\{j \mid k_j(h) \in g(j)\} \in \mathcal{U}$.

Also, $k_i(g)$ is the constant function with value $g(i)$.

For the semantics, we give the interpretation of a formula as a subset of $\mathcal{T}$.

- $[\![g \in h]\!] = \{q \mid \exists f \in h(q) \perp \in [\![k_q(g) = f]\!]\}$
- $[\![g = h]\!] = \{q \in \omega \mid \text{for all } f \in g(q) \perp \in [\![f \in k_q(h)]\!], \text{ and vice versa}\} \cup \{\perp \mid \forall f \in g(\perp) \perp \in [\![f \in h]\!] \text{ and vice versa, and } [\![g = h]\!] \in \mathcal{U}\}$
- $[\![\phi \wedge \psi]\!] = [\![\phi]\!] \cap [\![\psi]\!]$
- $[\![\phi \vee \psi]\!] = [\![\phi]\!] \cup [\![\psi]\!]$
- $[\![\phi \to \psi]\!] = [\![\psi]\!] \cup (\omega \backslash [\![\phi]\!]) \cup \{\perp \mid \perp \notin [\![\phi]\!] \wedge [\![\psi]\!] \cup (\omega \backslash [\![\phi]\!]) \in \mathcal{U}\}$
- $[\![\exists x\ \phi(x)]\!] = \{q \mid \exists h\ q \in [\![\phi(h)]\!]\}$
- $[\![\forall x\ \phi(x)]\!] = \{q \in \omega \mid \text{for all } h \perp \in [\![k_q(\phi)(h)]\!]\} \cup \{\perp \mid \text{for all } h \perp \in [\![\phi(h)]\!], \text{ and } [\![\forall x\ \phi(x)]\!] \in \mathcal{U}\}$

The actual models used in [13] differ from the above in that ω is partitioned into infinite subsets, and a copy of $\mathcal{U}$ is applied to each of those slices. For instance, ω could be partitioned into two subsets, the evens and the odds, and, in terms of an open set $\mathcal{O}$ on $\omega^{<\omega}$, if $\sigma \in \mathcal{O}$, then both $\{k \mid \sigma^\frown 2k \in \mathcal{O}\}$ and $\{k \mid \sigma^\frown 2k+1 \in \mathcal{O}\}$ must be in $\mathcal{U}$.

4.2.2 Topological Settling to External Sets

This topic was first developed to find a model of CZF_{Exp}, that is, CZF with the Subset Collection axiom replaced by Exponentiation, in which the Dedekind reals do not form a set [26]. Since that theory does suffice to show that the Cauchy reals form a set, this is a strong way of separating the Dedekind and Cauchy reals. Two aspects of such a construction quickly come to mind. One is that the Cauchy and Dedekind reals would have to be unequal; the most natural model for getting that is the topological model over the reals. The other is that Subset Collection must fail; the model for getting that which we saw above is settling to external sets. So the obvious candidate for the task at hand is the topological model over the reals with external settling, whatever that might mean. In fact, that's exactly what works.

Not surprisingly, the definition of this model, based on the reals with external settling, can be extended to external settling over any topological space [21]. The idea is that not only can a set be specified more by shrinking the open set you're taking as truth, as usual in topological models, you can also specialize to a point in the set. At that moment, the universe of sets you were looking at becomes the ground model. At the same time, a new universe of variable sets appears, based on the same space. We give the formal details.

Definition 2. *For a topological space T, a term is a set of the form $\{\langle J_i, \sigma_i\rangle \mid i \in I\} \cup \{\langle r_h, \sigma_h\rangle \mid h \in H\}$, where each σ is (inductively) a term, each J is an open set, each r is a member of T, and H and I are index sets.*

Definition 3. *For σ a term and $r \in T$, σ^r is defined inductively on the terms as $\{\langle T, \sigma_i^r\rangle \mid \langle J_i, \sigma_i\rangle \in \sigma \wedge r \in J_i\} \cup \{\langle T, \sigma_h^r\rangle \mid \langle r, \sigma_h\rangle \in \sigma\}$.*

Definition 4. *$J \Vdash \sigma = \tau$ iff for all $\langle J_i, \sigma_i\rangle \in \sigma$ $J \cap J_i \Vdash \sigma_i \in \tau$ and vice versa, and for all $r \in J$ $\sigma^r = \tau^r$*

$J \Vdash \sigma \in \tau$ iff for all $r \in J$ there is a $\langle J_i, \tau_i\rangle \in \tau$ and $J_r \subseteq J_i$ containing r such that $J_r \Vdash \sigma = \tau_i$

$J \Vdash \phi \wedge \psi$ iff $J \Vdash \phi$ and $J \Vdash \psi$

$J \Vdash \phi \vee \psi$ iff for all $r \in J$ there is a $J_r \subseteq J$ containing r such that $J_r \Vdash \phi$ or $J_r \Vdash \psi$

$J \Vdash \phi \to \psi$ iff for all $J' \subseteq J$ if $J' \Vdash \phi$ then $J' \Vdash \psi$, and, for all $r \in J$, there is a $J_r \subseteq J$ containing r such that, for all $K \subseteq J_r$, if $K \Vdash \phi^r$ then $K \Vdash \psi^r$

$J \Vdash \exists x\ \phi(x)$ iff for all $r \in J$ there is a $J_r \subseteq J$ containing r and a σ such that $J_r \Vdash \phi(\sigma)$

$J \Vdash \forall x\ \phi(x)$ iff for all σ $J \Vdash \phi(\sigma)$, and for all $r \in J$ there is a $J_r \subseteq J$ containing r such that for all σ $J_r \Vdash \phi^r(\sigma)$.

In this last definition, when $T = \mathbb{R}$, because of the homogeneity of the space, the case of $\to$ can be simplified to "for all $J' \subseteq J$ if $J' \Vdash \phi$ then $J' \Vdash \psi$, and, for all $r \in J$, if $\mathbb{R} \Vdash \phi^r$ then $\mathbb{R} \Vdash \psi^r$."

Theorem 12. *The model given by the semantics above satisfies Infinity, Pairing, Union, Extensionality, Set Induction, Bounded (i.e. Δ_0) Separation, and Collection. It also satisfies Eventual Power Set: for every X there is a C such that everything in C is a subset of X, and every subset of X is not unequal to everything in C. If T is locally connected then Exponentiation holds. If T is locally homogeneous then Full Separation holds.*

4.3 Topological Sideways Settling

There is a topological version of the sideways settling model from 3.3, iterated via immediate settling ([13], Theorem 5.7). Let $\mathcal{U}$ be a non-principal ultrafilter on ω; also, identify ω with $\omega \times \omega$. Let T be $\omega \cup \{\bot\}$. Take the discrete topology on ω; for neighborhoods of $\bot$, let $A \cup \{\bot\}$ be open exactly when each A_i is in $\mathcal{U}$, where A_i is the i^{th} slice of A, using here the identification of ω with $\omega \times \omega$. Take the submodel of the full model of those sets g that eventually settle down on slices: for some j and all $i \geq j$, the value of g is constant on the i^{th} section of ω.

5 Classical Outer Models

We have so far been using the idea of an outer model as something into which V embeds. At the same time, almost all of the constructions and examples we have seen can be done within V. For example, a full Heyting-valued model can be taken to be an extension of V by the inclusion of a generic; at the same time, just like with Boolean-valued models, the entire structure can be developed within V. It is at this point that the classical mathematician has the option to call Boolean-valued models forcing, pull in a generic G from outside of V, and work in the two-valued model $V[G]$. The constructivist does not have this option. Nevertheless, we can still avail ourselves of this two-valued outer model $V[G]$, to help us in building our constructive models.

An example of this we have seen already: the second example in 3.2.1, about Kripke settling to external sets, when we took eventual settling of a set in $V[G]$, G Cohen generic over V, to something in V.

In 6.1.2 below, we will see another example of the use of V-generics in combination with inner models.

For the rest of this section, we consider an application of generics over V to independence results around the Fan Theorem.

In sections 2.3 and 4.1, there were sketches of two models in which the Fan Theorem failed in various ways. Let's put this into context. The Fan Theorem says that every bar is uniform. The X-Fan Theorem, a.k.a. X-FAN, says that every X-bar is uniform, for any choice of a property X. We have already seen D-FAN, the Fan Theorem for decidable bars, and c-FAN, the Fan Theorem for c-bars. Extending the notion of a c-bar, a Π^0_1 bar is a bar which is definable as a Π^0_1 set, which can be understood as a set given by a Π^0_1 formula, or as an intersection of countably many decidable sets. Then Π^0_1-FAN states that every Π^0_1 bar is uniform. Over IZF, the following implications are trivial:

$$\text{Full FAN} \Rightarrow \Pi^0_1-\text{FAN} \Rightarrow c-\text{FAN} \Rightarrow D-\text{FAN} \Rightarrow \text{IZF}.$$

The natural question is whether any of those implications reverse. Some have long been known not to. For instance, the recursive realizability model shows that IZF does not prove D-FAN. As observed in [25], the construction from [8] shows that Π^0_1-FAN does not imply the full Fan Theorem. [3] shows that D-FAN does not prove c-FAN over a weak base theory. In [25] a technique is developed which shows all of these non-reversals, including that c-FAN does not imply Π^0_1-Fan. It is this technique, different from the two already discussed, that interests us here.

To develop some intuition, ask yourself, how could FAN fail? It is hard to see how FAN could fail without there being a specific counter-example. There are other instance in which a $\forall x \exists y$ statement fails without a counter-example, because there is just no uniform way of going from an x to a y. In the case of the Fan Theorem though, the y would be a finite level of the binary tree witnessing the uniformity of the bar. If there were such a y, it should be easy to find, by just going through the full binary tree level by level until the uniform bound is found. Instead, one can more readily imagine a counter-example, a set of nodes which is not uniform, but it still represents a bar, because one cannot find a path missing that set. How could that be the case? You can start at the root $\langle\rangle$; if only one of $\langle 0\rangle$ and $\langle 1\rangle$ is in the tree (i.e. is not in the bar), then your choice of successor is clear; you keep going until both children are in the tree – then what? If you happen to choose a child beyond which the entire tree dies, your goose is cooked. For X to be a non-uniform bar, there would have to be no way of predicting, when confronting a tree split, which if either of the two sub-trees will die. That's the cleverness of the Kleene tree: membership on this infinite tree is computable, yet it is not computable when a sub-tree will entirely die – there is no computable look-ahead mechanism. Hence if your universe consists only of computable objects, this tree provides a counter-example, actually a decidable counter-example, to the Fan Theorem.

It would be nice to get an example like the Kleene tree which is more amenable to manipulation, so that we can not only falsify D-FAN but also separate the varieties of FAN. What is the essence of the Kleene tree? From the point of view of computability, subtrees die out in seemingly random, unpredictable ways. The ultimate in set-theoretic unpredictability is forcing. A Kleene-like tree can be built generically. Let a forcing condition be an assignment to finitely many binary sequences of one of three labels: IN, meaning in the putative bar (as well as all of its descendants), OUT, meaning not in the bar although there's nothing to stop the bar from being uniform beneath it, and ∞, meaning the tree beneath that node is infinite. This infinitude beneath an ∞-labeled node is enforced in that at no time may a condition assign IN or OUT to all descendants at some fixed level; if all extensions of an ∞ node of length n are labeled, then at least one of those labels must be ∞. The generic G will be a tree with labels OUT sprinkled seemingly randomly beneath labels of ∞. Of course, one can easily find a path through this tree: at a node labeled ∞, take a child labeled ∞, which is guaranteed always to exist. What if we erase the labels ∞, and replace them with OUT? Notationally, this would be substituting G with $proj(G)$, the projection of G onto IN-OUT-valued nodes. This is exactly the information contained in a bar – not which nodes are part of infinite vs. finite sub-trees, but which nodes are in vs. out of the alleged bar. Not surprisingly, no infinite path through this tree will be computable from the tree.

This is of course just a start. One can erase the ∞ labels from the tree easily, but it is another matter to erase the ∞ information from the model. The difference between those OUT nodes rooting finite sub-trees and those rooting infinite ones must be obscured. At the same time, for the tree to remain decidable, we no longer have any choice about this labeling; once an IN-OUT decision has been made, we need to stick with it. The only solution seems to be to use non-standard integers. By taking the ultrapower of this model using a countably incomplete ultrafilter, we get an elementary extension $\mathcal{M}$ with non-standard integers. Now we can take the image of G, call it $i(G)$, in this ultrapower $\mathcal{M}$. There is nothing stopping us from changing finitely many ∞'s to OUTs in $i(G)$, even on standard nodes, and of course changing all of their standard descendants to OUTs also, as this does not affect $proj(G)$. What it does do is free us up to change all of the descendants on some non-standard level to INs. If we can do this coherently, then any branch that went through a node formerly labeled ∞ will indeed hit the alleged bar, albeit at a non-standard level, but no matter.

The way to do this coherently is to arrange all of the various ∞-OUT substitutions as nodes in a Kripke model, in the iterative style described in 4.1. The bottom node $\perp$ of this model is assigned the model $V[proj(G)]$, to use the

Kripke terminology from above. The immediate successors of $\bot$ are indexed by the IN-OUT labeled trees H in $\mathcal{M}$ which arise from $i(G)$ as described, of course as determined in $V[G]$, since a distinction has to be made between standard and non-standard nodes. Some of these trees H will end up being finite in the sense of $\mathcal{M}$; those yield terminal nodes in the Kripke model. Others will not, and the construction of the Kripke model continues on from there, with children of H's node being determined just the way the children of $\bot$ were. In the end, the tree $proj(G)$ will have no paths, because whatever a potential path might look like at some node, there will be some future node in which that path is already on a dead end.

As described, this model falsifies D-FAN. To get any of the other separations is just a matter of hiding the tree $proj(G)$ better. For instance, $proj(G)$ could be hidden as a c-tree (the complement of a c-bar), thereby falsifying c-FAN, and if this is done slyly enough, D-FAN will not be disturbed.

It should be mentioned that the model given does not make $\neg$D-FAN true, but merely make D-FAN not true. The reason is that the terminal nodes of the Kripke model are dense, so $\neg\neg$D-FAN is true. To get instead $\neg$D-FAN to hold, one could imagine iterating this construction from all of the terminal nodes. This might involve using settling as described above, although in a more intricate form. The partial order is more complicated, and we would constantly need to pull in generics from outside. This is left for future work.

6 Inner Models

Classical set theorists have identified and studied different kinds of inner models: L-like inner models for large cardinals, permutation sub-models of forcing extensions, HOD. Here we will examine two, L and permutation models. It would be interesting to see what happens and what could be done with HOD.

6.1 L: Constructibility Meets Constructivism

L is often called the universe of constructible sets, but this kind of constructibility has little to do with constructivism. Still, those notions are compatible, and it is an obvious question what happens with L constructively.

6.1.1 The Development of L

As first noticed by William Powell (in unpublished notes), the definition of L can remain essentially unchanged, mod the standard way of avoiding the successor-or-limit case split: $L_\alpha = \bigcup_{\beta\in\alpha} def(L_\beta)$, where $def(X)$ is the collection of definable subsets of X. (For a published version see [16], which also contains the rest of this sub-section.) One would then like to prove some of the basic theorems about L. It is at this point that the problems start.

The first goals would be to show that L is a model of IZF and that $L^L = L$ (a kind of absoluteness of L, from the ordinals). It plays a role just what the meta-theory is taken to be. Through most of this article, the meta-theory is for simplicity taken to be ZFC. If we were to do that when studying L, then one is left with classical L. Instead, what one wants is to develop L constructively, within IZF or something similar. The axioms at issue are the ones around Collection. Classically, over the other ZF axioms, Replacement, Collection, and

Reflection are all equivalent. The soft proofs that Reflection implies Collection, which implies Replacement, go through constructively. The reverse implications do not. So it can make a difference which version of IZF you're working in: IZF_{Rep} (with Replacement), IZF (with Collection), and IZF_{Ref} (with Reflection).

Reflection is very strong, and you can usually prove what you want to with Reflection easily. For instance, IZF_{Ref} shows that L satisfies IZF_{Ref} pretty easily. In contrast, in the cases at hand, Replacement seems just not to be enough. The issue is that if X is in L then there is some L_α that X is definable over; there just may not be a canonical such α (like the least). Let's examine the effect that has on a sample case, the proof of Replacement in L. Suppose that, in L, $\forall x \in X \ \exists! y \ \phi(x, y)$. One needs a bounding set, an L_α such that for all $x \in X$ there is some $\beta \in \alpha$ such that the witness y for x is definable over L_β. In order to use Replacement in the meta-theory to bound some such set of β's, one would have to pick out some unique such β, which is not clear is possible. It is easy to see though that Collection suffices: if $\forall x \in X \ \exists y \ \phi(x, y)$, then $\forall x \in X \ \exists \beta_x \ (\exists y \in def(L_{\beta_x}) \ \phi(x, y))$; using Collection one can then get a bounding set for the β_x's; of course, this bounding set would have to be turned into an ordinal α, by removing all of the non-ordinals and then taking the transitive closure; this α suffices to produce a bounding set L_α.

Checking the other IZF axioms in L is straightforward for most of them. The only challenge is Separation. The standard classical argument for Separation in L goes via Reflection, to which we do not have access. Instead, one has to argue inductively on formulas. The only difficult cases are the quantifiers. For $\exists$, consider $\phi(x) = \exists y \ \psi(x, y)$ and $X \in L$. The collection we want, $A = \{x \in X \mid \phi^L(x)\}$, is in any case a set in V, using Separation there. By the construction of A, $\forall x \in A \ \exists y \in L \ \psi^L(x, y)$. Using arguments similar to the proof of Collection in L, there is a bounding set $Y \in L$ for the y's necessary: $\forall x \in A \ \exists y \in Y \ \psi^L(x, y)$. By induction, we can use Separation to get $\{\langle x, y\rangle \in X \times Y \mid \psi^L(x, y)\}$ in L. The projection of that latter set onto its first components is definable, and hence in L.

The case of $\forall$ is even a bit trickier: $\phi(x) = \forall y \ \psi(x, y)$, $X \in L$. With $\exists$, the goal was clear: find a set big enough to include enough witnesses. With a universal statement, there are no witnesses. Instead, one considers the different subsets of X obtained by restricting the range of y: for each $C \in L$, let $A_C \in V$ be $\{x \in X \mid \forall y \in C \ \psi^L(x, y)\}$. (We use the notation A_L for the desired set $\{x \in X \mid \phi^L(x)\}$.) Notice that if $D \supseteq C$ then $A_D \subseteq A_C$. Consider, in V, $B = \{A_C \mid C \in L\}$. (Even though C here ranges over a class, B is a set, as a subcollection of the power set of X.) For each $b \in B$ there is an ordinal β such that $b = A_C$ for some $C \in L_\beta$. By Collection, there in an ordinal α including such a β for each $b \in B$. Moreover, since L_α is transitive, not only is any such C in L_α, but also $C \subseteq L_\alpha$. That means $A_{L_\alpha} \subseteq b$ for each $b \in B$. We would like to show that $A_{L_\alpha} = A_L$. For one direction, since $L \supseteq L_\alpha, A_L \subseteq A_{L_\alpha}$. In the other direction, suppose $x \in A_{L_\alpha}$. Let $y \in L$. Then $A_{\{y\}} \in B$. Since $A_{L_\alpha} \subseteq A_{\{y\}}$, $x \in A_{\{y\}}$, which means $\psi^L(x, y)$, as desired. Inductively, the set $E = \{\langle x, y\rangle \in X \times L_\alpha \mid \psi^L(x, y)\}$ is in L. From E, A_L is easily definable as $\{x \in X \mid \forall y \in L_\alpha \ \langle x, y\rangle \in E\}$.

Turning to the other goal, that $L^L = L$, the classical argument is that, inductively, α is definable over L_α as the collection of ordinals, so that ORD $\subseteq L$. Then definability over L_α is absolute, so that $L^L_\alpha = L_\alpha$, and you're done.

Constructively this falls apart immediately. It is not (always) the case that α is the set of ordinals in L_α. For example, consider something as simple as the two-node Kripke model, with nodes $\perp$ and $\top$. (We do not distinguish notationally in the following between sets in V and their canonical internalizations.) Let $1_\top$ be the set with no members at $\perp$ and 0 as a member at $\top$, so that $\top \Vdash 1_\top = 1$. Notice that $L_{1_\top} = \{0, 1_\top\}$. Let $\alpha = \omega \cup \{1_\top\}$. Then $L_\alpha = L_\omega \cup \{1_\top\}$. The sets definable over L_α include not only α, but also, for any natural numbers $k < n$, those sets x that look like k at $\perp$ (i.e. $\perp \Vdash y \in x$ iff $\perp \Vdash y < k$) and are equal to n at $\top$, which are all ordinals. By way of notation, $x^+ = x \cup \{x\}$. Then L_{α^+} includes all of those funny ordinals just described. So one can certainly define over L_{α^+} the set of ordinals, but that will be a strict superset of α^+. In this case, one can still get α^+ definably over L_{α^+}, but it should be clear that with a more elaborate example even that would not be possible. For instance, throw in some of those funny ordinals into α^+, the ones that look like k and then grow to n, for rather randomly chosen k's and n's, calling the result β. Definably over L_β are all of those funny ordinals, and there is no good way to pick out exactly which got put into β. So definably over L_β we can get a superset of β, but not β itself. So it is not clear that the ordinal β is in L.

That much being understood, there still is a very different construction to show that, under IZF, L in the sense of L is L. The reason is not that L contains all the ordinals, but rather that for every α there is an α^* in L such that $L_\alpha = L_{\alpha^*}$. This latter fact is shown inductively. One works in a set X large enough to include all the β^*'s for $\beta \in \alpha$. Within L, one cannot use α as a parameter to pick out exactly the β^*'s, because α may not be in L. Instead, one uses L_α as a parameter, which is in L. Take α^* to be the subset of X of all γ's such that $def(L_\gamma) \subseteq L_\alpha$, which is in L. By the choice of what goes in to α^*, $L_{\alpha^*} \subseteq L_\alpha$; since each β^* is in α^*, $L_{\alpha^*} \supseteq L_\alpha$.

6.1.2 Transferring Independence Results from V to L

Although the very basics of L carry over from ZF to IZF, the next level of results, AC and GCH, apparently do not. The problem seems to be that it is at best unclear how to do condensation arguments constructively. So ultimately it could be that the study of L is not so interesting constructively. If we do believe that, it would be nice to have at least some theorem or proof giving concrete evidence of such. One possibility is that there are constructions showing that it is not hard to get an arbitrary set into L by coding it into an ordinal (unpublished). The upshot is that constructively L might be a lot like V, even be V itself, regardless of how complicated V is. If it is so easy to get anything into L, even more so to get L to be V by expanding L, arguably that implies that there is no use in studying L for its own sake. Here we will sketch more modest examples tending in the same direction, translating classical independence results over V into constructive independence results over L [17].

The theories we will be considering are those around admissibility, or KP. Some gentle extensions of KP have been considered over the years. For instance, Π_2 Reflection implies the main KP axiom, Σ_1 Bounding. Also, Resolvability (see below for its statement) and Σ_1 Dependent Choice (as theories extending KP) each implies Π_2 Reflection. Over L, all of these theories are equivalent, but not in general: any implication that does not follow from what was just said does not hold, as can be demonstrated by forcing the appropriate reals and sets of reals (for details see [17]).

Our interest is of course in the constructive version of KP, namely IKP. This has not been studied much – the literature might well be limited to [2] and [17]. Perhaps this is because of CZF, which is a significant extension of IKP yet has the same proof-theoretic strength. Still, IKP is a perfectly coherent theory, and one can ask about independence results over it. Trivially independence results over KP are also independence results over IKP. Of interest here is to transfer the cited independence proofs over KP to independence proofs over IKP in L. The original proofs are based on generic reals and sets of such; the technique to effect the transfer is to code the generics as ordinals.

We describe the simplest example of such, that Resolvability does not imply Σ_1 DC. Resolvability is the axiom that the universe is the union of the range of a Δ_1 definable function on the ordinals. Although it is not usually described this way, the model of Resolvability + $\neg\Sigma_1$ DC in [17] is very well known, being the standard permutation model for the failure of the Axioim of Choice. That is, take countably many mutually generic Cohen reals $G_i, i \in \omega$, and the set G – not the sequence! – of these reals. If the ground model is L, then the permutation model $L(G)$ is the extension of L by each of the generics as well as the set G. The resolution is $L_\alpha[G]$ as α runs through the ordinals, and the failure of AC is actually a failure of Σ_1 DC.

The task is to get a model of IKP in which G and its members are in some sense reflected in ordinals, which then end up being in L. The model will be a Kripke model with underlying partial order $2^{<\omega}$. Given $\sigma \in 2^{<\omega}$, let 1_σ be the set which is forced to be 1 (i.e. $\{\emptyset\}$, as usual) by any τ extending or incompatible with σ, whereas any initial segment of σ does not force anything into 1_σ. A real (i.e. infinite binary sequence) R induces a branch B through the Kripke partial order. A first approximation to B is $\{1_\sigma \mid \sigma \text{ is an initial segment of } R\}$. Indeed, at any node σ along R that's exactly what we want. But if τ is incompatible with σ then at node τ the set just listed becomes $\{1\}$ and all information about B has been lost. So instead, at node τ, B is taken to consist of those sets 1_ρ such that, for all j between the lengths of τ and ρ, $\rho(j) = R(j)$. That is, the tail end of R, the part beyond the length of τ, is the path through the partial order (or rather the p.o.'s reflection in the ordinals) taken by B at τ. Including also 0 into B makes B an ordinal. Let B_i be the branch so induced by G_i. Then B_i is definable over L_{B_i}. Furthermore, letting β be the transitive closure of $\{B_i \mid i \in \omega\} \cup \{1_\sigma \mid \sigma \in 2^{<\omega}\}$, we get that $\{B_i \mid i \in \omega\}$ is definable over L_β. The ultimate model is gotten by iterating definability, starting with L_β, ω_1^{CK}-many times, which is enough to get IKP. Resolvability is, in essence, given by the very definition of the model, as an iteration along some ordinals of definability. The failure of Choice in $L[G]$ translates to the failure of Σ_1 DC here. It is easy to see that everything happening here stays in L.

The other models in [17] are of Π_2 Reflection plus the failures of Σ_1 DC and Resolvability, and of IKP plus the failure of Π_2 reflection, all in L of course.

6.2 Permutation Models

The first permutation model within constructive mathematics was Krol's [15], also described in [11] and modified in [28]. Here we will describe the permutation model from [19], because it has the expository advantage of being simpler.

The motivating question behind [19], from [7], was whether the Cauchy reals are Cauchy complete. Classically the reals can be construed many different ways, for instance as Cauchy sequences, or equivalence classes of such, or Dedekind cuts, which are all equivalent. But whatever you take them to be, the move from the rationals to the reals is a closure operator, in that it is idempotent: if you take Dedekind cuts of rationals, then a Dedekind cut of those things can be converted to an equivalent Dedekind cut of rationals. This all breaks down constructively. We have already seen (2.2) that the Dedekind and Cauchy reals might differ, and that a Cauchy sequence might have no modulus of convergence. It is not hard to imagine other related scenarios, such as a Cauchy sequence of Cauchy sequences which have no moduli of convergence, which itself is not equivalent to any Cauchy sequence of rationals. (One could try to diagonalize through the given sequence of sequences, but you don't know how far out to go in each one.) It was the purpose of [19] to go through all such possibilities, culminating in what seemed like the most difficult, the motivating question: a Cauchy sequence R_i, with modulus of convergence, of equivalence classes of pairs, each consisting of a Cauchy sequence r_{ij} of rationals and a modulus for it, which is inequivalent to any Cauchy sequence of rationals.

The limiting factor here is choosing a representative from each equivalence class. If one had that, then one could simply pick R_i within $\epsilon/2$ of the limit, and then within R_i pick r_{ij} within $\epsilon/2$ of its limit, to get a rational within ϵ of the ultimate limit. So the issue really is picking a representative of each equivalence class. This has a similar feel to the standard model of $\neg$AC, in which there is a set of reals with no canonical choice of member. So one is naturally led to think of permutation models. At the same time, there has to be more than that, since classically one can choose a representative from each equivalence class of Cauchy sequences. (For instance, use a case split as to whether the limit is rational or irrational; in the former case, a constant sequence will do, whereas in the latter, there is a unique rational of the form $1/n$ closest to the intended limit.) The solution, or at least a solution, is to take a permutation sub-model of a topological mode. Generically, take a Cauchy sequence G_i of Cauchy sequences g_{ij} (the latter consisting of rational numbers, and each having the same fixed modulus of convergence). Then take the sub-model of those sets X with support a finite set I of the indices i, meaning that arbitrary changes of the g_{ij}'s for $i \notin I$ which do not affect any limits do not change X. For instance, replacing a sequence $\langle g_{ij} \rangle_j$ with its equivalence class $[\langle g_{ij} \rangle_j]$ of Cauchy sequences with the same limit yields a set with support $\emptyset$. Similarly, replacing the sequence $\langle G_i \rangle_i$ with $\langle [G_i] \rangle_i$ produces a set with null support, which is exactly the example we want. No Cauchy sequence with finite support I can have a limit equal to the limit of $\langle [G_i] \rangle_i$, because the former depends only on $\{G_i \mid i \in I\}$ whereas the latter does not.

7 A Final Example

As a culmination of this material, we present a construction that uses many of the ideas and techniques developed here. The argument at the end is only sketched; a more rigorous development is to appear.

Although the purpose here is not that the result is of particular importance, but rather that the construction itself is hopefully appealing, still there can be no model without it being a model of something, and so we give the content background. Building on work of [4] and Fred Richman (personal communication), [6] defines a sequence (x_n) in a metric space to be *almost Cauchy* if for all $\epsilon > 0$ and strictly increasing $g : \mathbb{N} \to \mathbb{N}$ there is an $N \in \mathbb{N}$ such that, for all $n \geq N$, the diameter of $\{x_{g(n)}, x_{g(n)+1}, ..., x_{g(n+1)}\}$ is less than ϵ (meaning that, whenever $g(n) \leq i, j \leq g(n+1), d(x_i, x_j) < \epsilon$). Easily, every Cauchy sequence is almost Cauchy. Whether every almost Cauchy sequence is Cauchy is an interesting question, being implied by BD-N [4] but not following from set theory (in particular, IZF) alone [24]. Various similar formulations of this property are also discussed in [6], along with their implications among each other, some being equivalent to almost Cauchyness, others being merely implied by almost Cauchyness, all of which are mutually equivalent under Countable Choice. The obvious question is whether the equivalence holds in the absence of CC. The version we will discuss is the apparently weakest form, which considers not the entire sub-sequence indexed from $g(n)$ to $g(n+1)$, but rather just the endpoints: $\forall \epsilon > 0$, increasing $g : \mathbb{N} \to \mathbb{N}\ \exists N \in \mathbb{N}\ \forall n \geq N\ d(x_{g(n)}, x_{g(n+1)}) < \epsilon$, which in a display of unimaginativeness will be called here $<$-almost Cauchyness.

Theorem 13. *$<$-almost Cauchyness does not imply almost Cauchyness.*

Proof. We will provide a model with a $<$-almost Cauchy, not almost Cauchy counter-example (G_n^+), so called because it will be the non-negative part max(0, G_n) of a generic sequence (G_n). For (G_n^+) to be a counter-example, there must be some positive ϵ witnessing that. All positive numbers are roughly the same, so the witnessing ϵ will be 1. Also, the function g witnessing that (G_n^+) is not almost Cauchy for $\epsilon = 1$ must grow fast, because there would have to be enough room in the interval from $g(n)$ to $g(n+1)$ for the G^+-sequence to change significantly, without there being an h picking out indices in that interval that would contradict $<$-almost Cauchyness. It will turn out that the exponential function $g(n) = 2^n$ suffices. Then, for that choice of ϵ and g, to show that the diameter property of almost Cauchyness does not eventually hold ("$\exists N\ \forall n \geq N$"), it would seem at first that we must have unboundedly many examples n of its failure ("$\forall N\ \exists n \geq N$"), which would be difficult. It's easier to get one counter-example k, as long as that k is non-standard. The ultimate model will be a Kripke model based on a tree of height 1, with only standard integers at $\perp$ and including non-standard integers at the successor nodes. Then at $\perp$ there will be no N that works, because any N at $\perp$ must be standard, and there will be a counter-example larger than N at at least one of the successor nodes.

Thus far the discussion has been about falsifying almost Cauchyness. We must also validate $<$-almost Cauchyness. At this point, it's helpful to over-simplify matters and temporarily set our sights lower. Whereas $<$-almost Cauchyness is about all $\epsilon > 0$ and all increasing functions g, we will be concerned with only ϵ's and g's from the ground model V, as opposed to ones from the non-standard extension M. This helps because the counter-example k we want is non-standard and the ground model has less finesse when it comes to manipulating non-standard elements. In the end, though, we still have to account for elements from M. So the plan is to develop first an intermediate sequence (x_n), which will be $<$-almost Cauchy in the limited sense of ground model ϵ's and g's; the derivation of (G_n^+) will be exactly so that the ground model ϵ's and g's are the only ones that will be relevant. To summarize, our intermediate goal will be a sequence (x_n), upon which (G_n) and (G_n^+) will later be based, that satisfies the $<$-almost Cauchyness condition for ϵ and g from V, and seems to violate almost Cauchyness, in that for some i, j, and k non-standard with $2^k \leq i < j \leq 2^{k+1}$, $d(x_i, x_j) = 1$.

Using some standard model theory here, all of the things we need can be collected into a single type. Taking a, b, and k to be free variables, let's focus first on the violation of almost Cauchyness. Consider the set of axioms $2^k \leq a, a < b, b \leq 2^{k+1}$, and the infinite collection $b - a > 0, b - a > 1, b - a > 2$, etc. In any model of that type, we would define (x_n) as being 0 outside of the interval $[a, b]$, as increasing from 0 at x_a up to 1 at the mid-point $(a + b)/2$ of $[a, b]$ by equal-sized, infinitesimally small steps $2/(b - a)$, and then back down again to 0 at x_b by the same-sized steps. This gives the promised violation of almost Cauchyness, with i as a and j as $(a + b)/2$, with the added benefit that each step from x_n to x_{n+1} is only an infinitesimal change. The axioms written down are easily seen to be consistent by compactness.

Turning to validating $<$-almost Cauchyness, or at least the fragment of it promised, suppose some (increasing) h from V is such that h misses the interval (a, b) entirely:

$$\phi_h := \exists n \; h(n) \leq a \wedge b \leq h(n + 1).$$

Then h would be a confirming instance of $<$-almost Cauchyness, since $x_{h(m)} - x_{h(m+1)}$ will always be 0. And there will indeed be such instances, for example the function 2^n. This will not always be possible though, for example for the identity function. In that case, though, we have success for another reason: $x_m - x_{m+1}$ is always infinitesimal. More generally, for any standard natural number β (for "bound"), let

$$\psi_{h,\beta} := \forall n (\text{if } h(n) \text{ or } h(n+1) \text{is in the interval } (a, b) \text{ then } h(n + 1) - h(n) < \beta).$$

This case is another confirming instance, since then $x_{h(m)} - x_{h(m+1)}$ is always either 0 or infinitesimal, in any case less than every standard positive ϵ. For each $h \in V$ we would like to validate either ϕ_h or $\psi_{h,\beta}$ for some standard β. There is not always a natural choice between those options. For instance, consider the function that enumerates all the elements between 2^n and 2^{n+1} whenever n is even and omits that interval entirely when n is odd. Compare that with the same kind of function but interchanging the parity of n. One of those functions will fall on the ϕ side, the other on the ψ, and the choice of which is arbitrary.

Let Ty (for "type") be a maximal consistent set of formulas extending $\{2^k \leq a, a < b, b \leq 2^{k+1}, b - a > 0, b - a > 1, b - a > 2, ...\}$ with formulas of the form ϕ_h and $\psi_{h,\beta}$. Ty has size at most the continuum $\mathfrak{c}$. There are guaranteed to be realizers a, b, and k for Ty in a model M whenever every consistent set of formulas of size at most $\mathfrak{c}$ is realized in M. This property is called the $\mathfrak{c}^+$-saturation of M. It is a result of introductory model theory that an ultrapower is κ^+-saturated if the ultrafilter used to develop it is κ-regular, and that ZFC proves the existence of κ-regular ultrafilters for all κ (see e.g. [5], sec. 4.3, esp. 4.3.5 and 4.3.14). Pick such an M, a, b, and k.

We would like that for each $h \in V$ either ϕ_h or some $\psi_{h,\beta}$ is in Ty, whereas all we know so far is that Ty is maximal consistent. Toward this end, consider some such h. We will show that either ϕ_h or $\psi_{h,\beta}$ is consistent with Ty, and so by maximality will then be in Ty.

Say that n is *relevant* if $h(n)$ or $h(n + 1)$ is in the interval (a, b), the point being that only relevant n's affect the truth of either ϕ_h or $\psi_{h,\beta}$. If no n's are relevant then ϕ_h is true in M and hence consistent with Ty. Else consider the (necessarily non-empty) set of *gaps*, which are the numbers $h(n+1) - h(n)$ whenever n is relevant. If all gaps are standard, then, since the set of gaps is definable in M, they have a standard bound, say β. Immediately, $\psi_{h,\beta}$ is true in M and so consistent with Ty. Else there is a non-standard gap. Now suppose there is a non-standard gap with both endpoints (i.e. $h(n)$ and $h(n+1)$) in the interval $[a, b]$: $a \leq h(n)$ and $h(n+1) \leq b$. In this case, a and b could be re-interpreted to be $h(n)$ and $h(n+1)$ respectively. That interpretation would still satisfy Ty, and make ϕ_h true, hence consistent with Ty. If instead there is no such non-standard gap, then all of the non-standard gaps have to include one of the endpoints a or b: either $h(n) < a$ (and $a < h(n+1) < b$, in order for n to be relevant), or $b < h(n+1)$ (and similarly $a < h(n) < b$). For each of those two possibilities there is at most one such n. To summarize the current hypotheses, there is at least one and at most two non-standard gaps, each of which contains one of the endpoints a or b. We will show what to do when there are two. This will call for a two-step procedure. If instead there is only one, then only one of those steps need be done.

Toward this end, let $h(n) < a < h(n+1) < b$. If $h(n+1) - a$ is non-standard, then re-interpret b as $h(n+1)$ (and leave a as itself). Under this interpretation Ty remains true and ϕ_h becomes true, showing that ϕ_h is consistent with Ty. Else $h(n+1) - a$ is standard. Keep that in mind. Now consider the other non-standard gap, $a < h(m) < b < h(m+1)$. If $b - h(m)$ is non-standard, re-interpret a as $h(m)$. As above, Ty remains true and ϕ_h becomes true, so is consistent with Ty. Else $b - h(m)$ is standard, as is $h(n+1) - a$. In this case, re-interpret a as $h(n+1)$ and b as $h(m)$. Since the original distance between a and b was non-standard, and both were changed by only a standard amount, the distance between their re-interpretations is still non-standard. Hence all of Ty remains true. Furthermore, what had been the only non-standard gaps are no longer gaps, as m and n are no longer relevant. So all gaps (if any) are standard, and we can argue as above to get either ϕ_h or some $\psi_{h,\beta}$ consistent with Ty.

Now that we have a and b as desired, consider (x_n) as defined above. If we were to try to use (x_n) as our counter-example to almost Cauchyness, $<$-almost Cauchyness would fall flat on its face, because a and b, hence their mid-point, are readily definable from it. So we must hide a and b, by fuzzing (x_n) up. That calls for a topological model. The idea is to replace each value x_n for n between a and b with a small interval. If that's all we do, a and b will still be definable as the first and last places where the sequence is non-zero. So we consider such a space based on any pair i, j with $a \leq i < j \leq b$.

To be more precise, work for the moment in M. For any i and j with $a \leq i < j \leq b$, let $(x_n^{i,j})$ be the sequence which is 0 outside of $[i, j]$, starting at i grows by $2/(b-a)$ at each step until the mid-point of $[i, j]$, then shrinks by the same amount until it hits 0 again at j. Let $T^{i,j}$ consist of all sequences (y_n) which are 0 outside of $[i, j]$ and for which $\mid x_n^{i,j} - y_n \mid < 2/(b-a)$ for n in $[i, j]$. A basic open set is given by restricting each y_n to an open interval. Let $(G_n^{i,j})$ be the generic sequence, and $G_n^{i,j,+}$ be $\max(0, G_n^{i,j})$. For any n for which $x_n^{i,j} \geq 2/(b-a)$, there is no difference between $G_n^{i,j}$ and $G_n^{i,j,+}$. The importance of $(G_n^{i,j,+})$ is that in the other case there is an open set forcing $G_n^{i,j,+}$ to be 0.

Consider a Kripke model with bottom node $\perp$ and successor nodes indexed by i and j with $a \leq i < j \leq b$. Let G be a set in this Kripke frame which is a sequence (indexed by the natural numbers) of reals. At $\perp$, all of the natural numbers will be standard, and at such a standard n, G_n will be 0. At a successor node indexed by i and j, G will be $(G_n^{i,j,+})$. The ultimate model will be a version of $L[G]$. There is no harm in having taken V to be a model of $V = L$, so M satisfies the same. With that understanding, at a successor node i, j, the only difference between $L[G]$ and $M[(G_n^{i,j})]$ is the restriction of the generic to be non-negative. That can be described as follows. Suppose a term σ in the forcing language contains a member $\langle \mathcal{O}, \tau \rangle$ with some coordinate of $\mathcal{O}$ being an open interval (r, s) with $r < 0$. Then $\langle \mathcal{O}^{-\infty}, \tau \rangle$ is also in σ, where $\mathcal{O}^{-\infty}$ is $\mathcal{O}$ with all such intervals (r, s), $r < 0$, replaced by $(-\infty, s)$; furthermore, if $s < 0$ too, then (r, s) will be replaces by $(-\infty, 0)$; furthermore, this happens hereditarily. In short, open sets cannot be distinguished via their parts beneath 0. The universe at node i, j is a topological model, with truth values being open sets.

What is more critical here is to give the model at $\perp$. This is not a topological model. From the standpoint of the classical meta-theory, every sentence is either true at $\perp$ (and hence also at all successor nodes with top truth value) or not true at $\perp$. The universe and the semantics are defined by a simultaneous induction on the ordinals in V, using standard Kripke semantics and definability in L.

The set G still provides the counter-example to almost Cauchyness: if at $\perp$ for $\epsilon = 1$ and $g(n) = 2^n$ there were such an N, consider what happens at the node a, b. Why does $<$-almost Cauchyness hold? Suppose $\perp \Vdash g$ is a function from $\mathbb{N}$ to $\mathbb{N}$. At any successor node i, j, by the connectedness of the space, each value $g(n)$ is forced by the entire space $T^{i,j}$. If either i or j is changed by 1, call the new values i' and j'. The spaces $T^{i,j}$ and $T^{i',j'}$ overlap on a set which is open in both spaces. So both spaces force the same value for $g(n)$. That means that all successor nodes force the same value for $g(n)$. In particular, consider the node $i, i+1$ for some i. The open set $\mathcal{O}$ in which each component is $(-\infty, 0)$ also forces the same value for $g(n)$, all n, and also forces G to be the constant 0 sequence. So $\mathcal{O}$ forces the universe at node $i, i+1$ to be L, and g to be the image in M of its restriction to the standard natural numbers in V via its definition in L. Again, this g has the same values at all successor nodes. So everything forced at $\perp$ to be a function from $\mathbb{N}$ to $\mathbb{N}$ is given by a ground model function.

□

References

[1] Adam-Day, S., Yurii Khomskii, and Robert Lubarsky: to appear

[2] Avigad, J.: Interpreting Classical Theories in Constructive Ones. The Journal of Symbolic Logic 65, No. 4, pp. 1785-1812 (2000)

[3] Berger, J.: The logical strength of the uniform continuity theorem. In Beckman, Berger, Löwe, Tucker (eds.) Logical Approaches to Computational Barriers, Proceedings of CiE 2006, LNCS 3988, Springer, pp. 35 - 39 (2006)

[4] Berger, J., D. Bridges and E. Palmgren: Double Sequences, Almost Cauchyness and BD-N. Logic Journal of IGPL 20 (1), pp. 349–354 (2012)

[5] Chang, C.C. and H.J. Keisler: Model theory. Dover Books on Mathematics, Dover Publications (2013)

[6] Diener, H. and Robert Lubarsky: Notions of Cauchyness and Metastability. In Proceedings of LFCS 2018, Lecture Notes in Computer Science 10703 (Sergei N. Artemov and Anil Nerode, eds.), Springer, pp. 140-153 (2018)

[7] Escardo, M.H. and A.K. Simpson: A Universal Characterization of the Closed Euclidean Interval (Extended Abstract). Sixteenth Annual IEEE Symposium on Logic in Computer Science, pp.115-125 (2001)

[8] Fourman, Michael P. and J. Martin Hyland: Sheaf models for analysis. In Fourman, M.P., Mulvey, C.J., Scott, D.S. (eds.) Applications of Sheaves, Lecture Notes in Mathematics Vol. 753, pp. 280-301. Springer-Verlag, Berlin Heidelberg New York (1979)

[9] Gambino, Nicola: Heyting-valued interpretations for Constructive Set Theory. Annals of Pure and Applied Logic 137, pp. 164-188 (2006)

[10] Grayson, R.J.: Heyting-valued models for intuitionistic set theory. In Fourman, M.P., Mulvey, C.J., Scott, D.S. (eds.) Applications of Sheaves, Lecture Notes in Mathematics Vol. 753, pp. 402-414. Springer-Verlag, Berlin Heidelberg New York (1979)

[11] Grayson, R.J.: Concepts of general topology in constructive mathematics and in sheaves. Annals of Mathematical Logic 20, No. 1, pp. 1-41 (1981)

[12] Grayson, R.J.: Heyting-valued semantics. In Lolli, G., Longo, G., Marcja, A. (eds.) Logic Colloquium '82, Studies in Logic and the Foundations of Mathematics Vol. 112, pp. 181-208. North-Holland, Amsterdam New York Oxford (1984)

[13] Hendtlass, M. and R. Lubarsky: Separating fragments of WLEM, LPO, and MP. Journal of Symbolic Logic 81, No. 4, pp. 1315-1343 (2016), doi: 10.1017/jsl.2016.38

[14] Julian, W. and Fred Richman: A uniformly continuous function on [0,1] that is everywhere different from its infimum. Pacific Journal of Mathematics 111(2), pp. 333-340 (1984)

[15] Krol, M.D.: A Topological Model for Intuitionistic Analysis with Kripke's Scheme. Zeitschrift für mathematische Logik und Grundlagen der Mathematik 24, No. 25-30, pp. 427-436 (1978)

[16] Lubarsky, R.: Intuitionistic L. Logical Methods in Computer Science: The Nerode Conference (Crossley et al., eds.), Birkhauser, pp. 555-571 (1993)

[17] Lubarsky, R.: IKP and Friends. Journal of Symbolic Logic 67, pp. 1295-1322 (2002)

[18] Lubarsky, R.: Independence Results around Constructive ZF. Annals of Pure and Applied Logic 132, No. 2-3, pp. 209-225 (2005)

[19] Lubarsky, R.: On the Cauchy Completeness of the Constructive Cauchy Reals. Mathematical Logic Quarterly 53, No. 4-5, pp. 396-414 (2007)

[20] Lubarsky, R.: Geometric Spaces with No Points (and Addendum). Journal of Logic and Analysis 2, No. 6, http://logicandanalysis.org/, pp. 1-10 (2010), doi 10.4115/jla2010.2.6

[21] Lubarsky, R.: Topological Forcing Semantics with Settling. Proceedings of LFCS '09, Lecture Notes in Computer Science (Sergei N. Artemov and Anil Nerode, eds.), Springer (2009); also Annals of Pure and Applied Logic 163, pp. 820-830 (2012), doi 10.1016/j.apal.2011.09.014

[22] Lubarsky, R.: On the Failure of BD-N. Journal of Symbolic Logic 78, No. 1, pp. 39-56 (2013)

[23] Lubarsky, R.: Separating the Fan Theorem and Its Weakenings II. Proceedings of LFCS 2018 (Artemov and Nerode, eds.), LNCS 10703, Springer, pp. 242-255; also the Journal of Symbolic Logic, to appear

[24] Lubarsky, R. and Hannes Diener: Principles Weaker than BD-N. Journal of Symbolic Logic 78, No. 3, pp. 873-885 (2013)

[25] Lubarsky, R. and Hannes Diener: Separating the Fan Theorem and Its Weakenings. Proceedings of LFCS '13, Lecture Notes in Computer Science 7734 (Sergei N. Artemov and Anil Nerode, eds.), Springer, pp. 280-295 (2013); also Journal of Symbolic Logic 79, No. 3, pp. 792-813 (2014), doi: 10.1017/jsl.2014.9

[26] Lubarsky, R. and Michael Rathjen: On the Constructive Dedekind Reals. Proceedings of LFCS '07, Lecture Notes in Computer Science 4514 (Sergei N. Artemov and Anil Nerode, eds.), Springer, pp. 349-362 (2007); also Logic and Analysis 1, No. 2, pp. 131-152 (2008)

[27] Mostowski, A.W.: Proofs of non-deducibility in intuitionistic functional calculus. JSL 13, pp. 204-207 (1948)

[28] Scedrov, A.: Consistency and independence results in intuitionistic set theory. In Vol. 873: Constructive Mathematics, Proceedings of the Las Cruces meeting, 1980, Lecture Notes in Mathematics 873, F. Richman ed., Springer, pp. 54-86 (1981)

[29] Scott, D.S.: Extending the topological interpretation to intuitionistic analysis I. Compos. Math. 20, pp. 194-210 (1968)

[30] Scott, D.S.: Extending the topological interpretation to intuitionistic analysis II. In Myhill, J., Kino, A., Vesley, R.E. (eds.) Intuitionism and Proof Theory pp. 235-255. North-Holland, Amsterdam (1970)

[31] Stone, M.H.: Topological representations of distributive lattices and Brouwerian logics. Casopis Pro Pestvovani Mathematiky a Fysiki Cast Matematicka 67, pp. 1-25 (1937)

[32] Tarski, A.: Der Aussagenkalkül and die Topologie. Fundam. Math. 31, pp. 103-134 (1938)

[33] Troelstra, A.S., van Dalen, D.: Constructivism in Mathematics – An Introduction, Vol. II, Studies in Logic and the Foundations of Mathematics, Vol. 123. North-Holland, Amsterdam New York Oxford (1988)