A formalization of forcing and the unprovability of the continuum hypothesis

Jesse Michael Han[1]
Department of Mathematics, University of Pittsburgh
https://www.pitt.edu/~jmh288
jessemichaelhan@gmail.com

Floris van Doorn
Department of Mathematics, University of Pittsburgh
http://florisvandoorn.com/
fpvdoorn@gmail.com

Abstract

We describe a formalization of forcing using Boolean-valued models in the Lean 3 theorem prover, including the fundamental theorem of forcing and a deep embedding of first-order logic with a Boolean-valued soundness theorem. As an application of our framework, we specialize our construction to the Boolean algebra of regular opens of the Cantor space $2^{\omega_2 \times \omega}$ and formally verify the failure of the continuum hypothesis in the resulting model.

2012 ACM Subject Classification Theory of computation → Logic and verification; Theory of computation → Type theory; Software and its engineering → Formal methods

Keywords and phrases Interactive theorem proving, formal verification, set theory, forcing, independence proofs, continuum hypothesis, Boolean-valued models, Lean

Supplement Material https://github.com/flypitch/flypitch

Funding Both authors were supported by the Sloan Foundation, grant G-2018-10067.

Acknowledgements We thank the members of the Pitt-CMU Lean group, particularly Simon Hudon, Jeremy Avigad, Mario Carneiro, and Tom Hales for their feedback and suggestions; we are also grateful to Dana Scott and John Bell for their advice and correspondence.

Introduction

The continuum hypothesis (CH) states that there are no sets strictly larger than the countable natural numbers and strictly smaller than the uncountable real numbers. It was introduced by Cantor [7] in 1878 and was the very first problem on Hilbert's list of twenty-three outstanding problems in mathematics. Gödel [14] proved in 1938 that CH was consistent with ZFC, and later conjectured that CH is independent of ZFC, i.e. neither provable nor disprovable from the ZFC axioms. In 1963, Paul Cohen developed *forcing* [10, 11], which allowed him to prove the consistency of ¬CH, and therefore complete the independence proof. For this work, which marked the beginning of modern set theory, he was awarded a Fields medal—the only one to ever be awarded for a work in mathematical logic.

In this paper we discuss the formalization of a Boolean-valued model of set theory where the continuum hypothesis fails. The work we describe is part of the Flypitch project, which aims to formalize the independence of the continuum hypothesis. Our results mark a major milestone towards that goal.

Our formalization is written in the Lean 3 theorem prover. Lean is an interactive proof assistant under active development at Microsoft Research [12, 41]. It implements the

[1] Corresponding author.

Calculus of Inductive Constructions and has a similar metatheory to Coq, adding definitional proof irrelevance, quotient types, and a noncomputable choice principle. Our formalization makes as much use of the expressiveness of Lean's dependent type theory as possible, using constructions which are impossible or unwieldy to encode in HOL, much less ZF: Lean's ordinals and cardinals, which are defined as equivalence classes of well-ordered types, live one universe level up and play a crucial role in the forcing argument; the models of set theory we construct require as input an entire universe of types; our encoding of first-order logic uses parametrized inductive types to equate type-correctness with well-formedness, eliminating the need for separate well-formedness proofs.

The method of forcing with Boolean-valued models was developed by Solovay and Scott in '65-'66 [35, 38] as a simplification of Cohen's method. Some of these simplifications were incorporated by Shoenfield [40] into a general theory of forcing using partial orders, and it is in this form that forcing is usually practiced. While both approaches have essentially the same mathematical content (see e.g. [26, 23, 28]), there are several reasons why we chose Boolean-valued models for our formalization:

- **Modularity.** The theory of forcing with Boolean-valued models cleanly splits into several components (a general theory of Boolean-valued semantics for first-order logic, a library for calculations inside complete Boolean algebras, the construction of Boolean-valued models of set theory, and the specifics of the forcing argument itself) which could be formalized in parallel and then recombined.
- **Directness.** For the purposes of an independence proof, the Boolean-valued soundness theorem eliminates the need to produce a two-valued model. This approach also bypasses any requirement for the reflection theorem/Löwenheim-Skolem theorems, Mostowski collapse, countable transitive models, or genericity considerations for filters.
- **Novelty and reusability.** As far as we were able to tell, the Boolean-valued approach to forcing has never been formalized. Furthermore, while for the purposes of an independence proof, forcing with Boolean-valued models and forcing with countable transitive models accomplish the same thing, a general library for Boolean-valued semantics of a deeply embedded logic could be used for formal verification applications outside of set theory, e.g. to formalize the Boolean-valued semantics of stochastic λ-calculus [37, 4].
- **Amenability to structural induction.** As with Coq, Lean is able to encode extremely complex objects and reason about their specifications using inductive types. However, the user must be careful to choose the encoding so that properties they wish to reason about are accessible by structural induction, which is the most natural mode of reasoning in the proof assistant. After observing (1) that the Aczel-Werner encoding of ZFC as an inductive type is essentially a special case of the recursive *name* construction from forcing (c.f. Section 3), and (2) that the automatically-generated induction principle for that inductive type *is* $\in$-induction, it is easy to see that this encoding can be modified to produce a Boolean-valued model of set theory where, again, $\in$-induction comes for free.

We briefly outline the rest of the paper. In Section 1 we outline the method of Boolean-valued models and sketch the forcing argument. Section 2 discusses a deep embedding of first-order logic, including a proof system and the Boolean-valued soundness theorem. Section 3 discusses our construction of Boolean-valued models of set theory. Section 4 describes the formalization of the forcing argument and the construction of a suitable Boolean algebra for forcing $\neg$CH. Section 5 describes the formalization of some transfinite combinatorics. We conclude with a reflection on our formalization and an indication of future work.

1 Outline of the proof

ZFC is a collection of first-order sentences in the language of a single binary relation $\{\in\}$, used to axiomatize set theory. The continuum hypothesis can be written in this fashion as a first-order sentence CH. A proof of CH is a finite list of deductions starting from ZFC and ending at CH. The soundness theorem says that provability implies satisfiability, i.e. if $\mathsf{ZFC} \vdash \mathsf{CH}$, then CH interpreted in any model of ZFC is true. Taking the contrapositive, we can demonstrate the unprovability (equivalently, the consistency of the negation) of CH by exhibiting a single model where CH is not true.

A model of a first-order theory T in a language L is in particular a way of assigning true or false in a coherent way to sentences in L. Modulo provable equivalence, the sentences form a Boolean algebra and "coherent" means the assignment is a Boolean algebra homomorphism (so $\vee$ becomes join, $\forall$ becomes infimum, etc.) into $\mathbf{2} = \{\mathsf{true}, \mathsf{false}\}$. The soundness theorem ensures that this homomorphism v sends a proof $\phi \vdash \psi$ to an inequality $v(\phi) \leqslant v(\psi)$. $\mathbf{2}$ may be replaced by any complete Boolean algebra $\mathbb{B}$, where the top and bottom elements $\top, \bot$ take the place of true and false. It is straightforward to extend this analogy to a $\mathbb{B}$-valued semantics for first-order logic, and in this generality, the soundness theorem now says that for any such $\mathbb{B}$, if $\mathsf{ZFC} \vdash \mathsf{CH}$, then for any $\mathbb{B}$-valued structure where all the axioms of ZFC have truth-value $\top$, CH does also. Then as before, to demonstrate the consistency of the negation of CH it suffices to find just one $\mathbb{B}$ and a single $\mathbb{B}$-valued model where CH is not "true".

This is where forcing comes in. Given a universe V of set theory containing a Boolean algebra $\mathbb{B}$, one constructs in analogy to the cumulative hierarchy a new $\mathbb{B}$-valued universe $V^{\mathbb{B}}$ of set theory, where the powerset operation is replaced by taking functions into $\mathbb{B}$. Thus, the structure of $\mathbb{B}$ informs the decisions made by $V^{\mathbb{B}}$ about what subsets, hence functions, exist among the members of $V^{\mathbb{B}}$; the real challenge lies in selecting a suitable $\mathbb{B}$ and reasoning about how its structure affects the structure of $V^{\mathbb{B}}$. While $V^{\mathbb{B}}$ may vary wildly depending on the choice of $\mathbb{B}$, the original universe V always embeds into $V^{\mathbb{B}}$ via an operation $x \mapsto \check{x}$, and while the passage of x to $\check{x}$ may not always preserve its original properties, properties which are definable with only bounded quantification are preserved; in particular, $V^{\mathbb{B}}$ thinks $\check{\mathbb{N}}$ is $\mathbb{N}$.

To force the negation of the continuum hypothesis, we use the Boolean algebra $\mathbb{B} := \mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ of regular opens of the Cantor space $2^{\aleph_2 \times \mathbb{N}}$. For each $\nu \in \aleph_2$, we associate the $\mathbb{B}$-valued characteristic function $\chi_\nu : \mathbb{N} \to \mathbb{B}$ by $n \mapsto \{f \mid f(\nu, n) = 1\}$. This induces what $V^{\mathbb{B}}$ thinks is a new subset $\widetilde{\chi_\nu} \subseteq \mathbb{N}$, called a *Cohen real*, and furthermore, simultaneously performing this construction on all $\nu \in \aleph_2$ induces what $V^{\mathbb{B}}$ thinks is a function from $\check{\aleph}_2 \to \mathcal{P}(\mathbb{N})$. After showing that $V^{\mathbb{B}}$ thinks this function is injective, to finish the proof it suffices to show that $x \mapsto \check{x}$ preserves cardinal inequalities, as then we will have squeezed $\check{\aleph}_1$ properly between $\mathbb{N}$ and $\mathcal{P}(\mathbb{N})$. This is really the technical heart of the matter, and relies on a combinatorial property of $\mathbb{B}$ called the *countable chain condition* (CCC), the proof of which requires a detailed combinatorial analysis of the basis of the product topology for $2^{\aleph_2 \times \mathbb{N}}$; we handle this with a general result in transfinite combinatorics called the Δ-*system lemma*.

So far we have mentioned nothing about how this argument, which is wholly set-theoretic, is to be interpreted inside type theory. To do this, it was important to separate the mathematical content from the metamathematical content of the argument. While our objective is only to produce a model of ZFC satisfying certain properties, traditional presentations of forcing are careful to stay within the foundations of ZFC, emphasizing that all arguments may be performed internal to a model of ZFC, etc., and it is not immediately clear what parts of the argument use that set-theoretic foundation in an essential way and require modification in the passage to type theory. Our formalization clarifies some of these questions.

Sources Our strategy for constructing a Boolean-valued model in which CH fails is a synthesis of the proofs in the textbooks of Bell ([5], Chapter 2) and Manin ([27], Chapter 8). For the Δ-system lemma, we follow Kunen ([26], Chapters 1 and 5).

Viewing the formalization The code blocks in this paper were taken directly from our formalization, but for the sake of formatting and readability, we sometimes omit or modify universe levels, type ascriptions, and casts. We refer the interested reader to our repository,[2] which contains a guide on compiling and navigating the source files of the project.

2 First-order logic

The starting point for first-order logic is a *language* of relation and function symbols. We represent a language as a pair of $\mathbb{N}$-indexed families of types, each of which is to be thought of as the collection of relation (resp. function) symbols stratified by arity:

```
structure Language : Type (u+1) :=
(functions : ℕ → Type u) (relations : ℕ → Type u)
```

2.1 (Pre)terms, (pre)formulas

The main novelty of our implementation of first-order logic is the use of *partially applied* terms and formulas, encoded in a parametrized inductive type where the $\mathbb{N}$ parameter measures the difference between the arity and the number of applications. The benefit of this is that it is impossible to produce an ill-formed term or formula, because type-correctness is equivalent to well-formedness. This eliminates the need for separate well-formedness proofs.

Fix a language L. We define the type of **preterms** as follows:

```
inductive preterm : ℕ → Type u
| var {} : ∀ (k : ℕ), preterm 0
| func : ∀ {l : ℕ} (f : L.functions l), preterm l
| app : ∀ {l : ℕ} (t : preterm (l + 1)) (s : preterm 0), preterm l
```

We use de Bruijn indices to avoid variable shadowing. A member of `preterm n` is a partially applied term. If applied to `n` terms, it becomes a term. Every element of `preterm L 0` is a well-formed term. We use this encoding to avoid mutual or nested inductive types, since those are not too convenient to work with in Lean.

The type of **preformulas** is defined similarly:

```
inductive preformula : ℕ → Type u
| falsum {} : preformula 0 --  notation '⊥'
| equal (t₁ t₂ : term L) : preformula 0 -- notation '≃'
| rel {l : ℕ} (R : L.relations l) : preformula l
| apprel {l : ℕ} (f : preformula (l + 1)) (t : term L) : preformula l
| imp (f₁ f₂ : preformula 0) : preformula 0 -- notation ⟹
| all (f : preformula 0) : preformula 0 -- notation '∀'
-- ¬ f := f ⟹ ⊥, notation '~f'
-- ∃ f := ~ ∀' ~f, notation '∃' f
```

[2] https://github.com/flypitch/flypitch

A member of `preformula n` is a partially applied formula. If applied to `n` terms, it becomes a formula. Implication is the only binary connective. Since we use classical logic, we can define the other connectives from implication and falsum. Similarly, universal quantification is our only quantifier.

Our proof system is a natural deduction calculus, and all rules are motivated to work well with backwards-reasoning:

```
inductive prf : set (formula L) → formula L → Type u
| axm     {Γ A} (h : A ∈ Γ) : prf Γ A
| impI    {Γ} {A B} (h : prf (insert A Γ) B) : prf Γ (A ⟹ B)
| impE    {Γ} (A) {B} (h₁ : prf Γ (A ⟹ B)) (h₂ : prf Γ A) : prf Γ B
| falsumE {Γ} {A} (h : prf (insert ∼A Γ) ⊥) : prf Γ A
| allI    {Γ A} (h : prf Γ A) : prf Γ (∀' A)
| allE₂   {Γ} A t (h : prf Γ (∀' A)) : prf Γ (A[t // 0])
| ref     (Γ t) : prf Γ (t ≃ t)
| subst₂  {Γ} (s t f) (h₁ : prf Γ (s ≃ t)) (h₂ : prf Γ (f[s // 0])) :
          prf Γ (f[t // 0])
```

A member of `prf Γ A` is a proof tree encoding a derivation of A from Γ. Note that `prf` is `Type`- instead of `Prop`-valued, so different members of `prf Γ A` are not definitionally equal.

2.2 Completeness

As part of our formalization of first-order logic, we completed a verification of the Gödel completeness theorem. Although our present development of forcing did not require it, we anticipate that it will be useful later to e.g. prove the downward Löwenheim-Skolem theorem for extracting countable transitive models. Like soundness, it also serves as a proof-of-concept and stress-test of our chosen encoding of first-order logic.

For our formalization, we chose the Henkin-style approach of constructing a canonical term model. In order to perform the argument, which normally involves modifying the language "in place" to iteratively add new constant symbols, we had to adapt it to type theory. Since our languages are represented by pairs of indexed types instead of sets, we cannot really modify them in-place with new constant symbols. Instead, at each step of the construction, we must construct an entirely new language in which the previous one embeds, and in the limit we must compute a directed colimit of types instead of a union. This construction induces similar constructions on terms and formulas, and completing the argument requires reasoning with all of them. As a result of our design decisions, only a few arguments required anything more than straightforward case-analysis and structural induction. The final statement makes no restrictions on the cardinality of the language.

2.3 Boolean-valued semantics for first-order logic

A **complete Boolean algebra** is a type $\mathbb{B}$ equipped with the structure of a Boolean algebra and additionally operations Inf and Sup (which we write as $\bigsqcap$ and $\bigsqcup$) returning the infimum and supremum of an arbitrary collection of members of $\mathbb{B}$. We use $\sqcap, \sqcup, \implies, \top$, and $\bot$ to denote meet, join, material implication, and top/bottom elements. For more details on complete Boolean algebras, we refer the reader to the textbook of Halmos-Givant [13].

▸ **Definition 1.** Fix a language L and a complete Boolean algebra $\mathbb{B}$. A **$\mathbb{B}$-valued structure** is an instance of the following `structure`:

```
structure bStructure :=
(carrier : Type u)
(fun_map : ∀{n}, L.functions n → vector carrier n → carrier)
(rel_map : ∀{n}, L.relations n → vector carrier n → 𝔹)
(eq : carrier → carrier → 𝔹)
(eq_refl : ∀ x, eq x x = ⊤)
(eq_symm : ∀ x y, eq x y = eq y x)
(eq_trans : ∀{x} y {z}, eq x y ⊓ eq y z ≤ eq x z)
(fun_congr : ∀{n} (f : L.functions n) (x y : vector carrier n),
  ⨅(map2 eq x y) ≤ eq (fun_map f x) (fun_map f y))
(rel_congr : ∀{n} (R : L.relations n) (x y : vector carrier n),
  ⨅(map2 eq x y) ⊓ rel_map R x ≤ rel_map R y)
```

Above, "`⨅(map2 eq x y)`" means "the infimum of the list whose ith entry is `eq` applied to `x[i]` and `y[i]`".

Note that Boolean-valued equality is not really an equivalence relation, but "$\mathbb{B}$ thinks it is". One complication which then arises in Boolean-valued semantics is keeping track of the congruence lemmas for formulas. However, as part of the soundness theorem shows, once these extensionality proofs are provided for the basic symbols in the language, they extend by structural induction to all formulas.

2.4 The soundness theorem

A soundness theorem says that a proof tree may be replayed to produce an actual proof in the object of truth-values. When the object of truth-values is `Prop`, this says that a proof tree compiles to a proof term. When the object of truth-values is a Boolean algebra, this says that the proof tree becomes an internal implication from the interpretation of the context to the interpretation of the conclusion:

```
lemma boolean_soundness {Γ : set (formula L)} {A : formula L}
  (H : Γ ⊢ A) : ∀ M, (⨅γ ∈ Γ, M[γ]) ≤ M[A]
```

Of course, we also formalized the ordinary soundness theorem. As a result of our design decisions, the proofs of both the ordinary and Boolean-valued soundness theorems were straightforward structural inductions.

3 Constructing Boolean-valued models of set theory

Throughout this section, we fix a universe level u and a complete Boolean algebra $\mathbb{B}$: `Type u`.

In set theory (see e.g. Jech [23] or Bell [5]), Boolean-valued models are obtained by imitating the construction of the von Neumann cumulative hierarchy via a transfinite recursion where iterations of the powerset operation (taking functions into $\mathbf{2} = \{\text{true}, \text{false}\}$) are replaced by iterations of the "$\mathbb{B}$-valued powerset operation" (taking functions into $\mathbb{B}$).

Since this construction by transfinite recursion does not easily translate into type theory, our construction of Boolean-valued models of set theory is instead a variation on a well-known encoding originally due to Aczel [1, 3, 2]. This encoding was adapted by Werner [42] to encode ZFC into Coq, whose metatheory is close to that of Lean. Werner's construction was

implemented in Lean's `mathlib` by Carneiro [9]. In this approach, one takes a universe of types `Type u` as the starting point and then imitates the cumulative hierarchy by constructing the inductive type

```
inductive pSet : Type (u+1)
| mk (α : Type u) (A : α → pSet) : pSet
```

The Aczel-Werner encoding is closely related to the recursive definition of *names*, which is used in forcing to construct forcing extensions:

▸ **Definition 2.** Let P be a partial order (which one thinks of as a collection of forcing conditions). A P-*name* is a collection of pairs (y, p) where y is a P-name and $p : P$.

If P consists of only one element, then a P-name is specified by essentially the same information as a member of the inductive type `pSet` above. Conversely, specializing P to an arbitrary complete Boolean algebra $\mathbb{B}$, we generalize the definition of `pSet.mk` so that elements are recursively assigned Boolean truth-values:

```
inductive bSet (𝔹 : Type u) [complete_boolean_algebra 𝔹] : Type (u+1)
| mk (α : Type u) (A : α → bSet) (B : α → 𝔹) : bSet
```

Thus `bSet 𝔹` is the type of $\mathbb{B}$-names, and will be the underlying type of our Boolean-valued model of set theory. For convenience, if `x : bSet 𝔹` and `x := ⟨α, A, B⟩`, we put `x.type := α, x.func := A, x.bval := B`.

3.1 Boolean-valued equality and membership

In `pSet`, equivalence of sets is defined by structural recursion as follows: two sets x and y are equivalent if and only if for every $w \in x$, there exists a $w' \in y$ such that w is equivalent to w', and vice-versa. Analogously, by translating quantifiers and connectives into operations on $\mathbb{B}$, Boolean-valued equality is defined in the same way:

```
def bv_eq : ∀ (x y : bSet 𝔹), 𝔹
| ⟨α, A, B⟩ ⟨α', A', B'⟩ :=
              (⨅a : α, B a ⟹ ⨆a', B' a' ⊓ bv_eq (A a) (A' a')) ⊓
                (⨅a' : α', B' a' ⟹ ⨆a, B a ⊓ bv_eq (A a) (A' a'))
```

We abbreviate `bv_eq` with the infix operator $=^{\mathbb{B}}$. With equality in place, it is easy to define membership by translating "x is a member of y if and only if there exists a w indexed by the type of y such that $x = w$." As with equality, we denote $\mathbb{B}$-valued membership by $\in^{\mathbb{B}}$.

```
def mem : bSet 𝔹 → bSet 𝔹 → 𝔹
| a ⟨α' A' B'⟩ := ⨆a', B' a' ⊓ a =ᴮ A' a'
```

3.2 Automation and metaprogramming for reasoning in $\mathbb{B}$

As stressed by Scott [36], "A main point ... is that the well-known algebraic characterizations of [complete Heyting algebras] and [complete Boolean algebras] exactly mimic the rules of deduction in the respective logics." Indeed, that is really why the Boolean-valued soundness theorem is true. One thinks of the $\leqslant$ symbol in an inequality of Boolean truth-values as a turnstile in a proof state: the conjuncts on the left as a list of assumptions in context,

and the quantity on the right as the goal. For example, given `a b : 𝔹`, the identity $(a \Rightarrow b) \sqcap a \leqslant b$ could be proven by unfolding the definition of material implication, but it is really just modus ponens; similarly, given an indexed family `a : I → 𝔹`, the equivalence `(⨆i, a i ⩽ b) ↔ ∀ i, a i ⩽ b` is just ∃-elimination.

Difficulties arise when the statements to be proved become only slightly more complicated. Consider the following example, which should be "`by assumption`":

```
∀ a b c d e f g: 𝔹, (d ⊓ e) ⊓ (f ⊓ g ⊓ ((b ⊓ a) ⊓ c)) ⩽ a
```

or slightly less trivially, the following example where the goal is attainable by "just applying a hypothesis to an assumption"

```
∀ a b c d : 𝔹, (a ⟹ b) ⊓ c ⊓ (d ⊓ a) ⩽ b
```

There are three ways to deal with goals like these, which approximately describe the evolution of our approach. First, one can try using the basic lemmas in `mathlib`, using the simplifier to normalize expressions, and performing clever rewrites with the deduction theorem.[3] Second, one can take the LCF-style approach and expand the library of lemmas with increasingly sophisticated derived inference rules. Third, one can make the following observation:

▸ **Lemma 3** (Yoneda lemma for posets). *Let $(P, \leqslant)$ be a partially ordered set. Let $a\ b : P$. Then $a \leqslant b$ if and only if $\forall \Gamma : P, \Gamma \leqslant a \to \Gamma \leqslant b$.*

This is a consequence of the Yoneda lemma for partially ordered sets, and its proof is utterly trivial. However, one side of the equivalence is much easier for Lean to reason with. Take the example which should have been "`by assumption`". The following proof, in which the user navigates down the binary tree of nested ⊓s, will work:

```
example {a b c d e f g : 𝔹} : (d ⊓ e) ⊓ (f ⊓ g ⊓((b ⊓ a)⊓ c)) ⩽ a :=
by {apply inf_le_right_of_le, apply inf_le_right_of_le,
    apply inf_le_left_of_le, apply inf_le_right_of_le, refl}
```

But if we use the right-hand side of Lemma 3 instead, then after some preprocessing, `assumption` will literally work:

```
example {a b c d e f g : 𝔹} : (d ⊓ e) ⊓ (f ⊓ g ⊓((b ⊓ a)⊓ c)) ⩽ a :=
by {tidy_context, assumption}
-- `tidy_context` applies `poset_yoneda`, introduces a hypothesis `H`,
-- uses `simp` at H to convert ⊓s to ∧s, and automatically splits
/- Goal state before `assumption`:
[...]
H_right_right_left_left : Γ ⩽ b,
H_right_right_left_right : Γ ⩽ a
⊢ Γ ⩽ a -/
```

A key feature of Lean is that it is its own metalanguage, allowing for seamless in-line definitions of custom tactics. This feature was an invaluable asset, as it allowed the rapid development of a custom tactic library for simulating natural-deduction style proofs inside

[3] The deduction theorem in a Boolean algebra says that for all a, b and c, $a \sqcap b \leqslant c \iff a \leqslant b \Rightarrow c$.

$\mathbb{B}$ after applying Lemma 3. Boolean-valued versions of natural deduction rules like $\vee/\wedge$-elimination, instantiation of existentials, implication introduction, and even basic automation were easy to write. The result is that the user is able to pretend, with absolute rigor, that they are simply writing proofs in first-order logic while calculations in the complete Boolean algebra are being performed under the hood.

One use-case where automation is crucial is context-specialization. For example, suppose that after preprocessing with `poset_yoneda`, the goal is $\Gamma \leqslant$ `(a` $\Longrightarrow$ `b)`, and one would like to "introduce the implication", adding $\Gamma \leqslant$ `a` to the context and reducing the goal to $\Gamma \leqslant$ `b`. This is impossible as stated. Rather, the deduction theorem lets us rewrite the goal to $\Gamma \sqcap$ `a` $\leqslant$ `b`, and now we may add $\Gamma \sqcap$ `a` $\leqslant$ `a` to the context. So we may introduce the implication after all, but at the cost of specializing the context Γ to the smaller context Γ' `:=` $\Gamma \sqcap$ `a`. But now, in order for the user to continue the pretense that they are merely doing first-order logic, this change of variables must be propagated to the rest of the assumptions which may still be of the form $\Gamma \leqslant$ `_`—which is extremely tedious to do by hand, but easy to automate.

3.3 The fundamental theorem of forcing

The fundamental theorem of forcing for Boolean-valued models [17] states that for any complete Boolean algebra B, V^B is a Boolean-valued model of ZFC. Since, in type theory, a type universe `Type u` takes the place of the standard universe V, the analogous statement in our setting is that for every complete Boolean algebra $\mathbb{B}$, `bSet` $\mathbb{B}$ is a Boolean-valued model of ZFC.

Bell [5] gives an extremely detailed account of the verification of the ZFC axioms, and we faithfully followed his presentation for this part of the formalization. Most of it is routine. We describe some aspects of `bSet` $\mathbb{B}$ which are revealed by this verification.

Check-names

▸ **Definition 4.** From the definitions of `pSet` and `bSet`, one immediately sees that there is a canonical map `check : pSet → bSet` $\mathbb{B}$, defined by

```
def check : pSet → bSet 𝔹
| ⟨α,A⟩ := ⟨α, (λ a, check (A a)), λ a, ⊤⟩
```

We call members of the image of `check` *check-names*,[4] after the usual diacritic notation $\check{x}$ for `check (x : pSet)`. These are also known as *canonical names*, as they are the canonical representation of standard two-valued sets inside a Boolean-valued model of set theory.[5]

The axiom of infinity In `pSet`, ω is defined to be the collection of all finite von Neumann ordinals (via induction on $\mathbb{N}$), and (ω `: bSet` $\mathbb{B}$) is $\check{\omega}$. While it is easy to show $\check{\omega}$ satisfies the axiom of infinity

```
def axiom_of_infinity_spec (u : bSet 𝔹) : 𝔹 :=
  (∅∈ᴮ u) ⊓ (⨅i_x, ⨆i_y, (u.func i_x ∈ᴮ u.func i_y))
```

[4] This terminology is standard, c.f. [17, 28].

[5] We were pleased to discover Lean's support for custom notation allowed us to declare the Unicode modifier character `U+030C` (ˇ) as a postfix operator for `check`.

it can furthermore be shown to satisfy the universal property of ω, which says that ω is a subset of any set which contains $\varnothing$ and is closed under the successor operation $x \mapsto x \cup \{x\}$.

The axiom of powerset

▸ **Definition 5.** Fix a $\mathbb{B}$-valued set `x` $= \langle \alpha$, `A, b`$\rangle$. Let χ : $\alpha \to \mathbb{B}$ be a function. The subset of `x` associated to χ is a $\mathbb{B}$-valued set $\widetilde{\chi}$ defined as follows:

```
def set_of_indicator {x} (χ : x.type → 𝔹) := ⟨x.type, x.func, χ⟩
```

The **powerset** $\mathcal{P}(x)$ of x is defined to be the following $\mathbb{B}$-valued set, whose underlying type is the type of all functions `x.type` $\to \mathbb{B}$:

```
def bv_powerset (u : bSet 𝔹) : bSet 𝔹 :=
⟨u.type → 𝔹, (λ f, set_of_indicator f), (λ f, set_of_indicator f ⊆ᴮ u)⟩
```

The axiom of choice Following Bell, we verified Zorn's lemma, which is provably equivalent over ZF to the axiom of choice. As is the case with `pSet`, establishing the axiom of choice requires the use of a choice principle from the metatheory. This was the most involved part of our verification of the fundamental theorem of forcing, and relies on the technical tool of *mixtures*, which allow sequences of $\mathbb{B}$-valued sets to be "averaged" into new ones, and the *maximum principle*, which allows existentially quantified statements to be instantiated without changing their truth-value.

The smallness of $\mathbb{B}$ We end this section by remarking that the "smallness" (or more precisely, the fact that $\mathbb{B}$ lives in the same universe of types out of which `bSet` $\mathbb{B}$ is being built) is essential in making `bSet` $\mathbb{B}$ a model of ZFC. It is required for extracting the witness needed for the maximum principle, and is also required to even define the powerset operation, because the underlying type of the powerset is the function type of all maps into $\mathbb{B}$.

4 Forcing

4.1 Representing Lean's ordinals inside `pSet` and `bSet`

The treatment of ordinals in `mathlib` associates a class of ordinals to every type universe, defined as isomorphism classes of well-ordered types, and includes interfaces for both well-founded and transfinite recursion. Lean's ordinals may be represented inside `pSet` by defining a map `ordinal.mk : ordinal → pSet` via transfinite recursion; it is nothing more than the von Neumann definition of ordinals. In pseudocode,

```
def ordinal.mk : ordinal → pSet
| 0 := ∅
| succ ξ := pSet.succ (ordinal.mk ξ) -- (mk ξ ∪ {mk ξ})
| is_limit ξ := ⋃ η < ξ, (ordinal.mk η)
```

Composing by `check` (Definition 4) yields a map `check ∘ ordinal.mk : ordinal → bSet` $\mathbb{B}$. (We could just as well have defined `ordinal.mk′ : ordinal → bSet` $\mathbb{B}$ analogously to `ordinal.mk` without reference to `check`, such that `ordinal.mk′ = check ∘ ordinal.mk`; the point is that there is a link between the metatheory's notion of size and order with that of the forcing extension.)

Cardinals in Lean are defined separately from ordinals as bijective equivalence classes of types, but are canonically represented by ordinals which are not bijective with any predecessor. We let `aleph : ordinal → ordinal` index these representatives. For the rest of this section, unadorned alephs (e.g. "$\aleph_2$") will mean either an ordinal of the form `aleph ξ` or a choice of representative from the isomorphism class of well-ordered types, and checked alephs (e.g. "$\check{\aleph}_2$") will mean the `check ∘ ordinal.mk` of that ordinal.

4.2 The Cohen poset and the regular open algebra

Forcing with partial orders and forcing with complete Boolean algebras are related by the fact that every poset of forcing conditions can be embedded into a complete Boolean algebra as a dense suborder. This will be the case for our forcing argument: our Boolean algebra is the algebra of regular opens on $2^{\aleph_2 \times \mathbb{N}}$ (we identify this space with the subsets of $\aleph_2 \times \mathbb{N}$), and the poset of forcing condition embeds in this Boolean algebra as a dense suborder.

▸ **Definition 6.** The **Cohen poset** for adding $\aleph_2$-many Cohen reals is the collection of all finite partial functions $\aleph_2 \times \mathbb{N} \to \mathbf{2}$, ordered by reverse inclusion.

In the formalization, the Cohen poset is represented as a `structure` with three fields:

```
structure 𝒞 : Type :=
  (ins : finset (ℵ₂.type × ℕ))
  (out : finset (ℵ₂.type × ℕ))
  (H : ins ∩ out = ∅)
```

That is, we identify a finite partial function f with the triple $\langle$`f.ins, f.out, f.H`$\rangle$, where `f.ins` is the preimage of $\{1\}$, `f.out` is the preimage of $\{0\}$, and `f.H` ensures well-definedness. While f is usually defined as a finite partial function, we found that in practice f is really only needed to give a finite partial specification of a subset of $\aleph_2 \times \mathbb{N}$ (i.e. a finite set `f.ins` which *must* be in the subset, and a finite set `f.out` which *must not* be in the subset), and chose this representation to make that information immediately accessible.

▸ **Definition 7.** Let X be a topological space, and for any open set U, let $U^\perp$ denote the complement of the closure of U. The **regular open algebra** of a topological space X, written $\mathrm{RO}(X)$, is the collection of all open sets U such that $U = (U^\perp)^\perp$, equipped with the structure of a complete Boolean algebra, with $x \sqcap y := x \cap y$, $x \sqcup y := ((x \cup y)^\perp)^\perp$, $\neg x := x^\perp$, and $\bigsqcup x_i := ((\bigcup x_i)^\perp)^\perp$.

The Boolean algebra which we will use for forcing $\neg\mathsf{CH}$ is $\mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$. Unless stated otherwise, for the rest of this section, we put $\mathbb{B} := \mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$.

▸ **Definition 8.** We define the **canonical embedding** of the Cohen poset into $\mathbb{B}$ as follows:

```
def ι : 𝒞 → 𝔹 := λ p, {S | p.ins ⊆ S ∧ p.out ⊆ - S}
```

That is, we send each `c : 𝒞` to all the subsets which satisfy the specification given by `c`. This is a clopen set, hence regular. Crucially, this embedding is *dense*:

```
lemma 𝒞_dense {b : 𝔹} (H : ⊥ < b) : ∃ p : 𝒞, ι p ≤ b
```

Recalling that $\leqslant$ in $\mathbb{B}$ is subset-inclusion, we see that this is essentially because the image of $\iota : \mathcal{C} \to \mathbb{B}$ *is* the standard basis for the product topology. Our chosen encoding of the Cohen poset also made it easier to perform this identification when formalizing this proof.

4.3 Adding $\aleph_2$-many distinct Cohen reals

As we saw in Definition 5, for any $\mathbb{B}$-valued set x, characteristic functions into $\mathbb{B}$ from the underlying type of x determine $\mathbb{B}$-valued subsets of x. While the ingredients $\aleph_2$ and $\mathbb{N}$ for $\mathbb{B}$ are types and thus external to `bSet` $\mathbb{B}$, they are represented nonetheless inside `bSet` $\mathbb{B}$ by their check-names $\check{\aleph}_2$ and $\check{\mathbb{N}}$, and in fact $\aleph_2$ *is* $\check{\aleph_2}$ `.type` and $\mathbb{N}$ *is* $\check{\mathbb{N}}$`.type`. Given our specific choice of $\mathbb{B}$, this will allow us to construct an $\aleph_2$-indexed family of distinct subsets of $\check{\mathbb{N}}$, which we can then convert into an injective function from $\check{\aleph_2}$ to $\mathcal{P}(\mathbb{N})$, *inside* `bSet` $\mathbb{B}$.

▸ **Definition 9.** Let $\nu : \aleph_2$. For any $n : \mathbb{N}$, the collection of all subsets of $\aleph_2 \times \mathbb{N}$ which contain (ν, n) is a regular open of $2^{\aleph_2 \times \mathbb{N}}$, called the **principal open** $\mathbf{P}_{(\nu,n)}$ over (ν, n).

▸ **Definition 10.** Let $\nu : \aleph_2$. We associate to ν the $\mathbb{B}$-valued characteristic function $\chi_\nu : \mathbb{N} \to \mathbb{B}$ defined by $\chi_\nu(n) := \mathbf{P}_{(\nu,n)}$. In light of our previous observations, we see that each χ_ν induces a new $\mathbb{B}$-valued subset $\widetilde{\chi_\nu} \subseteq \check{\mathbb{N}}$. We call $\widetilde{\chi_\nu}$ a **Cohen real**.

This gives us an $\aleph_2$-indexed family of Cohen reals. Converting this data into an injective function from $\check{\aleph}_2$ to $\mathbb{N}$ inside `bSet` $\mathbb{B}$ requires some care. One must check that $\nu \mapsto \widetilde{\chi_\nu}$ is externally injective, and this is where the characterization of the Cohen poset as a dense subset of $\mathbb{B}$ (and moving back and forth between this representation and the definition as finite partial functions) comes in. Furthermore, one has to develop machinery similar to that for the powerset operation to convert an external injective function `x.type` $\to$ `bSet` $\mathbb{B}$ to a $\mathbb{B}$-valued set which `bSet` $\mathbb{B}$ thinks is a injective function, while maintaining conditions on the intended codomain. Our custom tactics and automation for reasoning inside $\mathbb{B}$ made this latter task significantly easier than it would have been otherwise. We refer the interested reader to our formalization for details.

4.4 Preservation of cardinal inequalities

So far, we have shown for $\mathbb{B} = \mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ that `bSet` $\mathbb{B}$ thinks $\check{\aleph}_2$ is smaller than $\mathcal{P}(\check{\mathbb{N}})$. Although Lean believes there is a strict inequality of cardinals $\aleph_0 < \aleph_1 < \aleph_2$, in general we can only deduce that their representations inside `bSet` $\mathbb{B}$ are subsets of each other: $\top \leqslant \check{\aleph}_0 \subseteq^{\mathbb{B}} \check{\aleph}_1 \subseteq^{\mathbb{B}} \check{\aleph}_2$. To finish negating CH, it suffices to show that `bSet` $\mathbb{B}$ thinks $\check{\aleph}_0$ is strictly smaller than $\check{\aleph}_1$, and that `bSet` $\mathbb{B}$ thinks $\check{\aleph}_1$ is a strictly smaller than $\check{\aleph}_2$. That is, for cardinals κ, we want that the passage from κ to $\check{\kappa}$ to preserve cardinal inequalities.

▸ **Definition 11.** For our purposes, "X is strictly smaller than Y" means "there exists no function `f` such that for every `y` $\in$ `Y`, there exists an `x` $\in$ `X` such that `(x,y)` $\in$ `f`". Thus, "`X` is strictly smaller than `Y`" translates to the Boolean truth-value

```
-(⨆f, (is_func f) ⊓ ⨅y, y ∈ᴮ Y ⟹ ⨆x, x ∈ᴮ X ⊓ (x, y) ∈ᴮ f).
```

We abbreviate this with "$X \prec Y$".

The condition on an arbitrary $\mathbb{B}$ which ensures the preservation of cardinal inequalities is the *countable chain condition*.

▸ **Definition 12.** We say that $\mathbb{B}$ has the **countable chain condition** (CCC) if every antichain $\mathcal{A} : I \to \mathbb{B}$ (i.e. an indexed collection of elements $\mathcal{A} := \{a_i\}$ such that whenever $i \neq j, a_i \sqcap a_j = \bot$) has a countable image.

We sketch the argument that CCC implies the preservation of cardinal inequalities. The proof is by contraposition. Let κ_1 and κ_2 be cardinals such that $\kappa_1 < \kappa_2$, and suppose that

$\check{\kappa_1}$ is not strictly smaller than $\check{\kappa_2}$. Then there exists some `f : bSet 𝔹` and some $\Gamma > \bot$ such that $\Gamma \leqslant$ `(is_func f) ⊓ ⨅y, y ∈ᴮ` $\check{\kappa_1}$ $\implies$ `⨆x, x ∈ᴮ` $\check{\kappa_2}$ `⊓ (x,y) ∈ᴮ f`. Then one can show:

```
lemma AE_of_check_larger_than_check :
∀ β < κ₂, ∃ η < κ₁, ⊥ < (is_func f) ⊓ (η̌, β ̌) ∈ᴮ f
```

The name of this lemma emphasizes that what has happened here is that, given this f and the assumption that it satisfes some ∀-∃ formula inside `bSet 𝔹`, we are able to extract, by virtue of $\check{\kappa_1}$ and $\check{\kappa_2}$ being check-names, a ∀-∃ statement in the *metatheory*. Using Lean's choice principle, we can then convert this ∀-∃ statement into a function $g : \kappa_2 \to \kappa_1$, such that for every β, `⊥ < (is_func f) ⊓ (g(β)̌ , β ̌) ∈ᴮ f`. Since $\kappa_2 > \kappa_1$, it follows from the infinite pigeonhole principle that there exists some $\eta < \kappa_1$ such that the $g^{-1}(\{\eta\})$ is uncountable. Define $\mathcal{A} : g^{-1}(\{\eta\}) \to \mathbb{B}$ by $\mathcal{A}(\beta) :=$ `(is_func f) ⊓ (g(β)̌ , β ̌) ∈ᴮ f`. This is an uncountable antichain because if $\beta_1 \neq \beta_2$, then the well-definedness part of `is_func f` ensures that, since $g(\beta_1) = g(\beta_2)$, the truth-value $\check{\beta_1} = f(g(\beta_1)) \neq^{\mathbb{B}} f(g(\beta_2)) = \check{\beta_2}$ is $\bot$.

Thus, conditional on showing that $\mathbb{B} = \mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ has the CCC, we now have that cardinal inequalities are preserved in `bSet 𝔹`. Combining this with the injection $\check{\aleph_2} \preceq \mathcal{P}(\mathbb{N})$, we obtain:

```
theorem neg_CH : ⊤ = (ℕ ≺ (ℵ₁)̌ ⊓ (ℵ₁)̌ ≺ (ℵ₂)̌ ⊓ (ℵ₂)̌ ≤ 𝒫(ℕ))
```

The arguments sketched in subsection 4.3 and subsection 4.4 form the heart of the forcing argument. Their proofs involve taking objects in `Type u` and `bSet 𝔹`, constructing corresponding objects on the other side, and reasoning about them in ordinary and $\mathbb{B}$-valued logic simultaneously to determine cardinalities in `bSet 𝔹`. We have omitted many details from our discussion, but of course, all the proofs have been formally verified.

4.5 The unprovability of CH

We conclude this section by briefly describing how the previous results may be converted into a formal proof of the unprovability of CH. We work in a conservative expansion ZFC′ of ZFC with an expanded language $L_{\mathsf{ZFC}'}$ with symbols for pairing, union, powerset, and ω. We define ZFC′ to be precisely the ZFC axioms which were verified in the fundamental theorem of forcing, along with specifications for the new function symbols. CH can then be written as a deeply-embedded $L_{\mathsf{ZFC}'}$ sentence (note the use of de Bruijn indices for variables)

```
def CH : sentence L_ZFC' := ¬ ∃' ∃' (ω ≺ &1) ⊓ (&1 ≺ &0) ⊓ (&0 ≤ 𝒫(ω))
```

where $\prec$ and $\preceq$ are abbreviations with the same meaning as in the previous section. Then proving `bSet 𝔹 ⊨ ZFC' + ¬CH` is a straightforward matter of checking that sentences are interpreted correctly as Boolean truth values which we have already proved to be $\top$. Applying the contrapositive of the Boolean-valued soundness theorem yields the result.

5 Transfinite combinatorics and the countable chain condition

What remains now is to prove that $\mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ has the CCC. There are several ways forward; we chose a very general proof using the Δ-system lemma to show more generally that the product of topological spaces satisfies the CCC if every finite subproduct does. Our proof follows Kunen [26].

5.1 The Δ-system lemma

▸ **Definition 13.** A family $(A_i)_i$ of sets is called a **Δ-system** (or a **sunflower** or **quasi-disjoint**) if there is a set r, called the **root** such that whenever $i \neq j$ we have $A_i \cap A_j = r$.

```
def is_delta_system {α ι : Type*} (A : ι → set α) :=
∃(root : set α), ∀{{x y}}, x ≠ y → A x ∩ A y = root
```

The Δ-system lemma states that if we have an uncountable family of finite sets, there is an uncounbtable subfamily which forms a Δ-system. In Lean this is formulated as follows. (`restrict A t` is the restriction of the collection `A` to `t`).

```
theorem delta_system_lemma_uncountable {α ι : Type*}
  (A : ι → set α) (h : cardinal.omega < mk ι)
  (h2A : ∀i, finite (A i)) : ∃(t : set ι),
  cardinal.omega < mk t ∧ is_delta_system (restrict A t)
```

This theorem follows from the following more general statement, taking $\kappa = \aleph_0$ and $\theta = \aleph_1$ (for cardinal numbers the operation `c ^< κ` or $c^{<\kappa}$ is the supremum of c^ρ for $\rho < \kappa$).

```
theorem delta_system_lemma {α ι : Type u} {κ θ : cardinal}
  (hκ : cardinal.omega ⩽ κ) (hκθ : κ < θ) (hθ : is_regular θ)
  (hθ_le : ∀(c < θ), c ^< κ < θ) (A : ι → set α)
  (hA : θ ⩽ mk ι) (h2A : ∀i, mk (A i) < κ) :
  ∃(t : set ι), mk t = θ ∧ is_delta_system (restrict A t)
```

We omit the proof, referring the interested reader to [26] or the formalization.

5.2 $\mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ has the countable chain condition

▸ **Definition 14.** We say that a topological space X satisfies the countable chain condition if every family of pairwise disjoint open sets is countable.

We first give a sufficient condition for a product of topological spaces to satisfy the countable chain condition.

▸ **Theorem 15.** *If we have a family $(X_i)_{i\in I}$ of topological spaces, then $\prod_{i\in I} X_i$ has the countable chain condition if for every finite $J \subseteq I$ the product $\prod_{i\in J} X_i$ has the countable chain condition.*

Proof. For the proof, suppose we had an uncountable family of pairwise disjoint open subsets U_k of $\prod_{i\in I} X_i$. By shrinking U_k, we may assume that each U_k is a basic open set of the form $\prod_{i\in F_k} U_{k,i} \times \prod_{i\notin F_k} X_i$ for some finite set F_k. Now the $(F_k)_k$ form a uncountable family of finite sets, so by the Δ-system lemma we know that there is an uncountable family K of indices such that $(F_k)_{k\in K}$ forms a Δ-system with root J. Now we can take the projections $\pi(U_k)$ onto $\prod_{i\in J} X_i$ for $k \in K$. We can show this forms an uncountable disjoint family of opens in $\prod_{i\in J} X_i$, contradicting the assumption. ◂

With this, the rest of the proof that $\mathbb{B} = \mathrm{RO}(2^{\aleph_2 \times \mathbb{N}})$ has the CCC is easy: since every finite product 2^J is a finite topological space, and so satisfies the CCC, it follows that the space $2^{\aleph_2 \times \mathbb{N}}$ satisfies the CCC. Also, if a topological space X satisfies the CCC then the algebra of regular opens satisfies the CCC, since every antichain of regular opens forms a family of disjoint open sets. Thus, we have shown:

```
theorem 𝔹_CCC : CCC (regular_opens (set(ℵ₂.type × ℕ)))
```

6 Related work

First-order logic, soundness, and completeness There are many existing formalizations of first-order logic. Shankar [39] used a deep embedding of first-order logic to formalize incompleteness theorems. Harrison gives a deeply-embedded implementation of first-order logic in HOL Light [18] and a proof-search style account of the completeness theorem in [19]. Margetson [33] and Schlichtkrull [34] use the same argument for the completeness theorem in Isabelle/HOL, while Berghofer [6] (in Isabelle) and Ilik [22] (in Coq) use canonical term models.

Set theory and forcing Set theory is a common target for formalization. Notably, a large body of formalized set theory has been completed in Isabelle/ZF, led by Paulson and his collaborators [32, 29, 30]. Most relevantly, this includes a formalization of the relative consistency of the axiom of choice with ZF [31]. Building on this, Gunther, Pagano, and Terraf have begun formalizing the basic ingredients of forcing [15, 16], taking the more conventional approach of generic extensions of countable transitive models.

Our tactic library for Boolean-valued logic was inspired by work of Hudon [21] on Unit-B, using similar techniques to embed a proof language for temporal logic [20]. It was pointed out to the authors that a trick similar to Lemma 3 had also been successfully applied in the Metamath library [8].

The work we have described in this paper relies heavily on Lean's `mathlib`. In particular, the extensive `set_theory` and `ordinal` libraries contained nearly everything we needed (including a treatment of cofinalities for the Δ-system lemma), with missing parts easily accessible through existing lemmas. These libraries were originally developed by Carneiro [9], in part to show that Lean proves the existence of infinitely many inaccessible cardinals.

7 Conclusions and future work

Reflections on the proof

As our formalization has shown, for the purposes of a consistency proof, one can perform forcing entirely outside of the set-theoretic foundations in which forcing is usually presented. There is no need to work inside an ambient model of set theory, or to even have a ground model of set theory over which one constructs a forcing extension. Instead, the recursive *name* construction applied to a universe of types is key. The type universe, with its classical two-valued logic and its own notion of ordinals, takes the place of the standard universe of sets. These external ordinals are then represented in the internal ordinals of the forcing extension by indexing the construction of von Neumann ordinals. With a clever choice of forcing conditions $\mathbb{B}$, this representation of ordinals will preserve cardinal inequalities and force an uncountable set beneath $\mathcal{P}(\mathbb{N})$.

In particular, `pSet`, being only another special case of the construction which produces `bSet` $\mathbb{B}$, is no longer a prerequisite for working with `bSet` $\mathbb{B}$, but merely a convenient tool for organizing the check-names—this is the only role it played in the proof. The check-names themselves were actually not necessary either: as we remarked, the canonical map `ordinal` $\to$ `bSet` $\mathbb{B}$ can be defined without reference to them. However, since in all of our sources, `pSet` additionally played the role of the universe of types, and an interface for it was readily available in `mathlib`, we started our formalization by following the usual arguments, implementing these simplifications as we became aware of them.

Lessons learned

- Originally, we thought set-theoretic arguments involving transfinite/ordinal induction, which are ubiquitous, would be difficult to implement. In practice, Lean's tools for well-founded recursion and the comprehensive treatment of ordinals in `mathlib` made the implementation of such arguments painless.
- Definitions and lemmas should be stated as generally as possible. This maximizes reusability, minimizes redundancy, and by exposing only the information required to complete the proof, improves the performance of automation.
- One should invest early in domain-specific automation. The formalization of the fundamental theorem was completed using only the first two strategies outlined in subsection 3.2; the calculations, while tedious, were recorded in our sources and it seemed easier to follow them. If we had followed through on the observations around Lemma 3 and developed the custom tactic library earlier, we would have saved a significant amount of time.

Towards a formal proof of the independence of the continuum hypothesis

The work we have described in this paper was undertaken as part of the Flypitch project, which aims to produce a formal proof of the independence of the continuum hypothesis. As such, the obvious next goal is a formalization of the consistency of CH. Although it would be possible to do this using Boolean-valued models, we intend to develop the infrastructure necessary to support a proof by forcing with generic extensions, as well as Gödel's original proof by way of analyzing the constructible universe L.

Although our work includes a formal proof of the unprovability of a version of CH from a version of the ZFC axioms in a conservative extension of the language of ZFC, verifying this after completing the forcing argument (as in subsection 4.5) is easy. What is more interesting is formalizing the equivalence of various common formulations of ZFC and CH, so that a skeptical user may verify that their preferred version of CH is unprovable from their preferred version of ZFC. This would require formalizations of the conservativity of commonly-used extensions of ZFC, and of the equivalence of the various ways to say that one set is strictly smaller than another. The proof of the completeness theorem already required formalizing nontrivial conservativity statements, which shows that our framework is well-equipped to support such results.

Although the stated goal of our project is to achieve a formal proof of the independence of the continuum hypothesis, we also intend to develop reusable libraries for set theory and mathematical logic. We have completed a formalization of forcing, but are nowhere near completing a library which a set theorist could use to verify their research. Just as, more than 50 years ago, Cohen's proof marked the beginning of modern research in set theory, a formal proof of the independence of the continuum hypothesis will only mark the beginning of an integration of formal methods into modern research in set theory. This will require robust interfaces for handling the diverse range of forcing arguments and for reasoning about the consistency strengths of various extensions of ZFC, so that—to paraphrase Kanamori [24, 25]—deeply-embedded notions of truth and relative consistency become matters of routine manipulation as in algebra. Our work demonstrates that such tasks are well within the scope of modern interactive theorem provers.

8 References

References

1 Peter Aczel. The type theoretic interpretation of constructive set theory. In *Logic Colloquium*, volume 77, pages 55–66, 1978.

2 Peter Aczel. The type theoretic interpretation of constructive set theory: choice principles. In *Studies in Logic and the Foundations of Mathematics*, volume 110, pages 1–40. Elsevier, 1982.

3 Peter Aczel. The type theoretic interpretation of constructive set theory: inductive definitions. In *Studies in Logic and the Foundations of Mathematics*, volume 114, pages 17–49. Elsevier, 1986.

4 Giorgio Bacci, Robert Furber, Dexter Kozen, Radu Mardare, Prakash Panangaden, and Dana Scott. Boolean-valued semantics for the stochastic λ-calculus. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 669–678. ACM, 2018.

5 John L Bell. *Set theory: Boolean-valued models and independence proofs*, volume 47. Oxford University Press, 2011.

6 Stefan Berghofer. First-order logic according to Fitting. *Archive of Formal Proofs*, August 2007. `http://isa-afp.org/entries/FOL-Fitting.html`, Formal proof development.

7 Georg Cantor. Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik*, 84:242–258, 1878.

8 Mario Carneiro. Natural deduction in the Metamath proof explorer. `http://us.metamath.org/mpeuni/mmnatded.html`, 2014. Slides (`http://us.metamath.org/ocat/natded.pdf`).

9 Mario Carneiro. The type theory of Lean. In preparation (`https://github.com/digama0/lean-type-theory/releases`), 2019.

10 Paul J Cohen. The independence of the continuum hypothesis. *Proceedings of the National Academy of Sciences*, 50(6):1143–1148, 1964.

11 Paul J Cohen. The independence of the continuum hypothesis, II. *Proceedings of the National Academy of Sciences*, 51(1):105, 1964.

12 Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. The Lean theorem prover (system description). In Amy P. Felty and Aart Middeldorp, editors, *Automated Deduction - CADE-25 - 25th International Conference on Automated Deduction, Berlin, Germany, August 1-7, 2015, Proceedings*, volume 9195 of *Lecture Notes in Computer Science*, pages 378–388. Springer, 2015. URL: `https://doi.org/10.1007/978-3-319-21401-6_26`, `doi:10.1007/978-3-319-21401-6\_26`.

13 Steven Givant and Paul Halmos. *Introduction to Boolean algebras*. Springer Science & Business Media, 2008.

14 Kurt Gödel. The consistency of the axiom of choice and of the generalized continuum-hypothesis. *Proceedings of the National Academy of Sciences*, 24(12):556–557, 1938.

15 Emmanuel Gunther, Miguel Pagano, and Pedro Sánchez Terraf. First steps towards a formalization of forcing. *CoRR*, abs/1807.05174, 2018. URL: `http://arxiv.org/abs/1807.05174`, `arXiv:1807.05174`.

16 Emmanuel Gunther, Miguel Pagano, and Pedro Sánchez Terraf. Mechanization of separation in generic extensions. *CoRR*, abs/1901.03313, 2019. URL: `http://arxiv.org/abs/1901.03313`, `arXiv:1901.03313`.

17 Joel David Hamkins and Daniel Evan Seabold. Well-founded boolean ultrapowers as large cardinal embeddings. *arXiv preprint arXiv:1206.6075*, 2012.

18 John Harrison. Formalizing basic first order model theory. In Jim Grundy and Malcolm C. Newey, editors, *Theorem Proving in Higher Order Logics, 11th International Conference, TPHOLs'98, Canberra, Australia, September 27 - October 1, 1998, Proceedings*, volume 1479 of *Lecture Notes in Computer Science*, pages 153–170. Springer, 1998. URL: `https://doi.org/10.1007/BFb0055135`, `doi:10.1007/BFb0055135`.

19 John Harrison. *Handbook of Practical Logic and Automated Reasoning*. Cambridge University Press, 2009.

20 Simon Hudon. Temporal logic in Unit-B, 2018. https://github.com/unitb/temporal-logic.

21 Simon Hudon, Thai Son Hoang, and Jonathan S. Ostroff. The Unit-B method: refinement guided by progress concerns. *Software & Systems Modeling*, 15:1091–1116, 2015.

22 Danko Ilik. *Constructive completeness proofs and delimited control*. PhD thesis, Ecole Polytechnique X, 2010.

23 Thomas Jech. *Set theory*. Springer Science & Business Media, 2013.

24 Akihiro Kanamori. The mathematical development of set theory from Cantor to Cohen. *Bulletin of Symbolic Logic*, 2(1):1–71, 1996.

25 Akihiro Kanamori. *The Higher Infinite: Large Cardinals in Set Theory from their beginnings*. Springer Science & Business Media, 2008.

26 Kenneth Kunen. *Set theory*, volume 102 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam-New York, 1980.

27 Yu I Manin. *A course in mathematical logic for mathematicians*, volume 53. Springer Science & Business Media, 2009.

28 Justin Tatch Moore. The method of forcing. *arXiv preprint arXiv:1902.03235*, 2019.

29 Lawrence C. Paulson. Set Theory for Verification: I. From foundations to functions. *J. Autom. Reasoning*, 11(3):353–389, 1993. URL: https://doi.org/10.1007/BF00881873, doi:10.1007/BF00881873.

30 Lawrence C. Paulson. The reflection theorem: A study in meta-theoretic reasoning. In Andrei Voronkov, editor, *Automated Deduction - CADE-18, 18th International Conference on Automated Deduction, Copenhagen, Denmark, July 27-30, 2002, Proceedings*, volume 2392 of *Lecture Notes in Computer Science*, pages 377–391. Springer, 2002. URL: https://doi.org/10.1007/3-540-45620-1_31, doi:10.1007/3-540-45620-1_31.

31 Lawrence C. Paulson. The relative consistency of the axiom of choice - mechanized using isabelle/zf. In Arnold Beckmann, Costas Dimitracopoulos, and Benedikt Löwe, editors, *Logic and Theory of Algorithms, 4th Conference on Computability in Europe, CiE 2008, Athens, Greece, June 15-20, 2008, Proceedings*, volume 5028 of *Lecture Notes in Computer Science*, pages 486–490. Springer, 2008. URL: https://doi.org/10.1007/978-3-540-69407-6_52, doi:10.1007/978-3-540-69407-6_52.

32 Lawrence C. Paulson and Krzysztof Grabczewski. Mechanizing set theory. *J. Autom. Reasoning*, 17(3):291–323, 1996. URL: https://doi.org/10.1007/BF00283132, doi:10.1007/BF00283132.

33 Tom Ridge and James Margetson. A mechanically verified, sound and complete theorem prover for first order logic. In Joe Hurd and Thomas F. Melham, editors, *Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005, Oxford, UK, August 22-25, 2005, Proceedings*, volume 3603 of *Lecture Notes in Computer Science*, pages 294–309. Springer, 2005. URL: https://doi.org/10.1007/11541868_19, doi:10.1007/11541868_19.

34 Anders Schlichtkrull. *Formalization of logic in the Isabelle proof assistant*. PhD thesis, Technical University of Denmark, 2018.

35 Dana Scott. A proof of the independence of the continuum hypothesis. *Theory of Computing Systems*, 1(2):89–111, 1967.

36 Dana Scott. The algebraic intepretation of quantifiers: intuitionistic and classical. *Andrzej Mostowski and Foundational Studies*, pages 289–312, 2008.

37 Dana Scott. Stochastic λ-calculi. *Journal of Applied Logic*, 12(3):369–376, 2014.

38 Dana Scott and Robert Solovay. Boolean algebras and forcing. Unpublished manuscript, 1967.

39 Natarajan Shankar. *Metamathematics, machines and Gödel's proof*, volume 38. Cambridge University Press, 1997.

40 Joseph R Shoenfield. Unramified forcing. In *Axiomatic set theory*, volume 13, pages 357–381. AMS Providence, RI, 1971.

41 Sebastian Ullrich. Lean 4: A guided preview. https://leanprover.github.io/talks/vu2019.pdf. Slides.

821 **42** Benjamin Werner. Sets in types, types in sets. In *International Symposium on Theoretical*
822 *Aspects of Computer Software*, pages 530–546. Springer, 1997.