Proofs for Higher-Order SMT and Beyond

Chad E. Brown[1], Mikoláš Janota[1], and Cezary Kaliszyk[2]

[1] Czech Technical University in Prague
Czech Institute of Informatics, Robotics and Cybernetics
Prague, Czech Republic
Mikolas.Janota@cvut.cz
[2] University of Innsbruck
Innsbruck, Austria
cezary.kaliszyk@uibk.ac.at

Abstract

to do

1 Introduction

A preliminary proposal for SMT-LIB Version 3.0 was recently published online [1]. According to this proposal, there are plans to extend SMT in serious ways, essentially bringing an expressive power somewhere between Church's simple type theory [11] (by including arrow types) and the Calculus of Inductive Constructions [22, 8, 23] (by including dependent types and inductively defined types). In addition, a working group on SMT proofs was announced [7] with the goal of developing a standard for "producing independently checkable proofs." Of course, having a standard notion of proof for SMT3 will require clarifying the intended semantics of SMT3 so that there is precision about what sets of formulas should be unsatisfiable (so there might be a "proof" of inconsistency) or satisfiable (so there might be a "model"). We consider the possibility of using higher-order set theory via the well-known Werner-Aczel semantics of Calculus of Inductive Constructions to provide both a clear semantics and a notion of checkable proof that is likely to be sufficient for SMT3 as well as possible future extensions. We also give examples to demonstrate the feasibility of the approach.

2 Models and Proofs in General

In the best case scenario a logic provides a clear definition of propositions, a rigorous definition of when a proposition is provable and a class of interpretations with a satisfaction relation. A proposition is considered valid if it is true in every interpretation in the class. The logic satisfies soundness and completeness if provability coincides with validity. The most well-known case is classical first-order logic with any number of proof systems and interpretations given by Tarski-style semantics.

Church's simple type theory provides another example of such a logic. In Church's original paper [11] there is a clear definition of types, terms (some of which are propositions) and a Hilbert style proof system. Henkin [15] later gave a notion of semantics for which a completeness result could be proven. (Technically Henkin's interpretations were not all sound with respect to Church's functional extensionality axiom, but this was corrected by Andrews [4].) An equality-based version of Church's simple type theory with a Hilbert style proof system and a notion of interpretation (called *general models*) following the Henkin-Andrews approach is presented in [3]. Furthermore in [3] one can find proofs of soundness, completeness and the usual results

associated with first-order logic such as the Lowenheim-Skolem Theorem and the Compactness Theorem.

For more serious extensions of Church's simple type theory – such as the Calculus of Inductive Constructions – there does not seem to be an effort to create a Henkin-Andrews notion of interpretation for which one could prove soundness and completeness. Instead research into semantics for type theories has tended to go in the direction of category theory [20, 17] and the most interesting interpretations are not classical.

In terms of soundness alone, there is one well-known set theoretic interpretation of type theories like the Calculus of Inductive Constructions. The interpretation is classical, extensional and satisfies proof irrelevance.[1] It was described by Werner [27] and Aczel [2] with more details found in the works of Werner, Lee and Barras [21, 5]. In this model, the universe of propositions is interpreted as a two element set – one of which is empty (having no proofs) representing "false" and the other being a singleton (having one proof) representing "true." Being a two element set makes it essentially the same as the interpretation of the type of booleans, as seems to be the intended treatment of propositions as booleans in SMT. Types are interpreted as sets (including the empty set) which live in some universe closed under various set theoretic operations. Coq is a well-known proof assistant based on the Calculus of Inductive Constructions (CIC) and each type universe is closed under the formation of (dependent) function types and inductively defined types. The Werner-Aczel style of interpretation would interpret each of Coq's universes as a set U closed under the corresponding set-theoretic operations (e.g., if A and B are in the set U, then the set B^A of functions is in the set U).

An alternative to attempting to obtain a Henkin-Andrews style semantics for which soundness and completeness can be proven is to simply take the standard set theoretic semantics *but* allow the model of the underlying set theory to change. That is, instead of defining a proposition as valid if it is true in every standard set theoretic interpretation, one could define it as being valid if it is true in every standard set theoretic interpretation living in a model of, say, first-order ZFC. Validity would then become recursively enumerable again and we clearly have a complete proof system (given by any proof system for first-order ZFC). We explore this possibility in this paper, except we use higher-order Tarski Grothendieck (HOTG) as described in [10] instead of first-order ZFC. The reason for using higher-order instead of first-order is to make the theory finitely axiomatizable. (We still obtain complete calculi via Henkin-Andrews semantics.) The reason for using Tarski Grothendieck instead of Zermelo Fraenkel is to ensure we have sufficient set theoretic universes for interpreting the type theoretic universes of CIC. For more information, a longer discussion is in the unpublished paper [9], from which some of the material form this article was taken.

3 Higher-Order Set Theory with Proof Terms

We begin by giving a formulation of simple type theory with proof terms, which we then extend to include set theory. The types are simple types and the terms are simply typed λ-terms in the style of Church [11]. The proof system is a natural deduction system [24] that admits proof terms in the usual Curry-Howard-de Bruijn style [16, 14, 25]. We additionally include constants and axioms for Tarski-Grothendieck style set theory [26] similar to the formulation described in [10].

We have two base types ι (sets) and o (propositions). All other types are function types of the form $(\alpha\beta)$ of functions from α to β. Such function types are often written as $(\alpha \to \beta)$.

[1] Proof irrelevance means all proofs of a given proposition are equal.

$$\frac{}{\Gamma \vdash \mathsf{Known}_s : s}\ s \in \mathcal{A} \qquad \frac{}{\Gamma \vdash u : s}\ u : s \in \Gamma \qquad \frac{\Gamma \vdash \mathcal{D} : s}{\Gamma \vdash \mathcal{D} : t}\ s{\approx}t \qquad \frac{\Gamma, u : s \vdash \mathcal{D} : t}{\Gamma \vdash (\lambda u : s.\mathcal{D}) : s \to t}$$

$$\frac{\Gamma \vdash \mathcal{D} : s \to t \qquad \Gamma \vdash \mathcal{E} : s}{\Gamma \vdash (\mathcal{D}\mathcal{E}) : t} \qquad \frac{\Gamma \vdash \mathcal{D} : s}{\Gamma \vdash (\lambda x.\mathcal{D}) : \forall x.s}\ x \in \mathcal{V}_\alpha \setminus \mathcal{F}\Gamma \qquad \frac{\Gamma \vdash \mathcal{D} : \forall x.s}{\Gamma \vdash (\mathcal{D}t) : s_t^x}\ x \in \mathcal{V}_\alpha, t \in \Lambda_\alpha$$

$$\frac{}{\Gamma \vdash \mathsf{Ext}_{\alpha,\beta} : (\forall fg : \alpha\beta.(\forall x : \alpha.fx = gx) \to f = g)}\ f, g \text{ DISTINCT}$$

Figure 1: Natural Deduction Calculus with Proof Terms

When parentheses are omitted they should be replaced to the right, e.g., $\iota\iota o$ is the type $(\iota(\iota o))$.

Let $\mathcal{V}_\alpha$ be the set of variables of type α and $\mathcal{S}_\alpha$ be a set of constants of type α. Assume we have countably many variables at each type. We now define a family $(\Lambda_\alpha)_\alpha$ of terms recursively, where $s \in \Lambda_\alpha$ means s is a term of type α.

- (Variables) If $x \in \mathcal{V}_\alpha$, then $x \in \Lambda_\alpha$.
- (Constants) If $c \in \mathcal{S}_\alpha$, then $c \in \Lambda_\alpha$.
- (Application) If $s \in \Lambda_{\alpha\beta}$ and $t \in \Lambda_\alpha$, then $(st) \in \Lambda_\beta$.
- (Abstraction) If $x \in \mathcal{V}_\alpha$ and $t \in \Lambda_\beta$, then $(\lambda x.t) \in \Lambda_{\alpha\beta}$.
- (Implication) If $s \in \Lambda_o$ and $t \in \Lambda_o$, then $(s \to t) \in \Lambda_o$.
- (Universal Quantification) If $x \in \mathcal{V}_\alpha$ and $t \in \Lambda_o$, then $(\forall x.t) \in \Lambda_o$.

We use common conventions for omitting parentheses and abbreviating multiple binders. Propositions are terms in Λ_o. The set $\mathcal{F}(s)$ for free variables of a term is defined as usual as is the notion of capture avoiding substitution, denoted s_t^x. We consider two terms to be equal if they are the same up to α-conversion (renaming of bound variables). The notion of $\beta\eta$-conversion, denoted $s{\approx}t$, is also defined in the usual way.

Given a family of constants $\mathcal{S}$ and a set of propositions $\mathcal{A}$, we can give a notion of provability via a natural deduction system. We give such a system, annotated with proof terms, in Figure 1. It is straightforward to write a proof checker for such a calculus. Indeed it uses the same ideas as the earliest proof checker, AUTOMATH [13, 14], dating back to 1968. A particular proof checker is included in software supporting the Proofgold cryptocurrency.[2] In our examples we will create proofs checkable by the Proofgold proof checker. We will also compare the performance of the checker distributed with the Proofgold Core software to a much faster alternative implementation due to the third author.

Our primary use case is where $\mathcal{S}$ is a collection of set theoretic constants (either primitive or defined) and $\mathcal{A}$ is a set of propositions that are either axioms of set theory or follow from those axioms. The particular set theory we have in mind is a form of higher-order Tarski-Grothendieck (HOTG). The primitive constants are those listed in [10], with the exception that we only take the choice operator ε_ι at type ι, rather than at every type. Specifically we have $\varepsilon_\iota : (\iota o)\iota$, $\mathsf{In} : \iota\iota o$, $\mathsf{Empty} : \iota$, $\mathsf{Union} : \iota\iota$, $\mathsf{Power} : \iota\iota$, $\mathsf{Repl} : \iota(\iota\iota)\iota$ and $\mathsf{UnivOf} : \iota\iota$. The axioms we have in

[2] https://prfgld.github.io

mind are those given in [10], again with the exception that we only take a choice axiom at type ι. The axioms are sufficient to ensure the logic is classical and extensional. As a consequence the proof system is sound and complete with respect to Henkin-Andrews semantics.

From now on we will write set theoretic propositions in the usual mathematical style, with the understanding that this can be (and is) fully formalized. For example, $\forall xAB.x \in A \wedge A \subseteq B \to x \in B$ corresponds to $\forall x.\forall A.\forall B.\mathsf{and}\ (\mathsf{In}\ x\ A)\ (\mathsf{Subq}\ A\ B) \to \mathsf{In}\ x\ B$ where $x, A, B \in \mathcal{V}_\iota$, In is primitive, $\mathsf{and} : ooo$ is defined in the usual Russell-Prawitz style as $\lambda q.\lambda r.\forall p.(q \to r \to p) \to p$, and $\mathsf{Subq} : \iota\iota o$ is defined as $\lambda A.\lambda B.\forall x.\mathsf{In}\ x\ A \to \mathsf{In}\ x\ B$.

4 Translating to Set Theory

The Werner-Aczel interpretation of the Calculus of Inductive Constructions (CiC) is described elsewhere [27, 2, 21, 5]. For our purposes we simply write $M, N, A, B, D, \ldots$ for terms of CiC and assume we have a partial function which may assign a set $\mathcal{T}_\varphi M$ to M, given an assignment φ for (at least) the variables in M. We assume that for well-typed terms M depending on variables $x_1 : A_1, \ldots x_n : A_n$, $\mathcal{T}_\varphi M$ is defined whenever $\varphi x_i \in \mathcal{T}_\varphi A_i$ for $i \in \{1, \ldots, n\}$. We furthermore assume the values satisfy the expected properties. For example, if M has type A, then $\mathcal{T}_\varphi M \in \mathcal{T}_\varphi A$. In particular, if M is a proposition (has type Prop), then $\mathcal{T}_\varphi M \in 2$, where 2 is $\{0, 1\}$. Here, 0 is the empty set and 1 is $\{0\}$. The value 0 is also assigned to every proof. That is, if M is a proposition with proof D, then $\mathcal{T}_\varphi M$ is 1 and $\mathcal{T}_\varphi D$ is 0 (for appropriate assignments φ).

Intuitively $\mathcal{T}$ maps from a type theory (CiC) to the language of mathematics. However, our intention is to use $\mathcal{T}$ to map from CiC to the formal set theory in Section 3. This provides both a semantics to CiC and a different (stronger) notion of proof term, the notion of proof from Section 3. While there is no proof of proof irrelevance in CiC, there is a proof of its translation via $\mathcal{T}$.

As a starting point for translating SMT to set theory, let us consider sorts and terms in SMT to be corresponding terms in CiC. In that case, $\mathcal{T}$ already provides a method of translating SMT sorts and terms to sets. If we simply consider SMT propositions to be terms of type boolean, then we can also translate SMT propositions to sets (each provably a member of 2) – a set which is "true" if 0 is a member of it and "false" otherwise. However, the SMT propositions will correspond more closely to the set theoretic propositions if we use $\mathcal{T}$ to define a mapping $\mathcal{T}^p$ sending SMT propositions to set theoretic propositions. For example, $\mathcal{T}^p_\varphi(\neg P)$ should be $\neg\mathcal{T}^p_\varphi(P)$, $\mathcal{T}^p_\varphi(\forall x : A.P)$ should be $\forall x : \iota.x \in \mathcal{T}_\varphi(A) \to \mathcal{T}^p_{\varphi, x\mapsto x}(P)$ and $\mathcal{T}^p_\varphi(s = t)$ should be $\mathcal{T}_\varphi(s) = \mathcal{T}_\varphi(t)$. If no other case applies, $\mathcal{T}^p_\varphi(P)$ is taken to be $0 \in \mathcal{T}_\varphi(P)$.

It is an oversimplification to consider SMT sorts to be CiC types. Some SMT sorts have a special intended meaning. For example, the SMT sort Int of integers should be interpreted as the set of integers, i.e., we should have $\mathcal{T}_\varphi(\mathsf{Int}) = \omega \cup \{-n | n \in \omega\}$, where $-n$ is defined appropriately. In the examples in this paper we will only use the SMT sorts for booleans, integers and arrays. Hence we assume $\mathcal{T}_\varphi(\mathsf{Bool}) = 2$ and $\mathcal{T}_\varphi(\mathsf{Int})$ is the set of integers. The interpretation of arrays is restricted but not fixed by the specification (see Page 39 of [6]), and we will handle this in a special way shown in the next section.

Suppose an SMT problem is given by a set of declarations of sorts $\sigma_1, \ldots, \sigma_n$, typed constants $u_1 : \alpha_1, \ldots, u_m : \alpha_m$ and assertions $P_1, \ldots, P_k$. Let U be a fixed Grothendieck universe, i.e., a set provably satisfying the properties of ZFC. We can translate the SMT problem to the set theoretic proposition

$$\forall \sigma_1 \cdots \sigma_n \in U.\forall u_1 \in \mathcal{T}_{\varphi_1}(\alpha_1) \cdots \forall u_m \in \mathcal{T}_{\varphi_1}(\alpha_m).\mathcal{T}^p_{\varphi_2}(P_1) \to \cdots \to \mathcal{T}^p_{\varphi_2}(P_k) \to \bot$$

where φ_1 takes each α_i to a corresponding variable of type ι (a "set") which we also call α_i and φ_2 extends φ_1 by also taking each u_j to a corresponding variable of type ι (a "set") which we also call u_j. Note that the set theoretic proposition corresponding to the SMT problem is *provable* if the SMT problem is *unsatisfiable*. As a consequence, if the negation of the set theoretic proposition is provable, then the SMT problem must be satisfiable. It is also possible that neither the set theoretic proposition nor its negation is provable.

5 Examples

We now consider a few examples. In each case we will show the result of translating the problem to a formal set theory and note there is either a formal proof of the set theoretic proposition or a formal proof of its negation. We briefly describe the proofs in each case. To make definitions and construct proofs in the formal set theory we will use the Megalodon system[3] (the successor to the Egal system [10]). Megalodon can also produce Proofgold proof terms presented in a simple to parse prefix notation.[4] While the Proofgold checker can be used for type checking and proof checking the data, we claim that it is straightforward to implement an independent proof checker and we additionally check the proofs with a faster reimplementation of the checker. We allow ourselves to freely use previous definitions or previously proven results (if they have been previously proven in Megalodon and published in Proofgold documents). That is, we do not need the proof term to contain a justification back to the axioms of set theory, but only back to previously proven results.

5.1 Induction

As a first simple example we consider induction on the natural numbers. Here the natural numbers are considered as a predicate over the sort Int.

In SMT2 format we can assert induction fails (which should be unsatisfiable) by giving a predicate p which holds for 0 and is closed under successor but does not hold for all integers $n \geq 0$. Here is the SMT2 specification:

```
(declare-fun p (Int) Bool)
(assert (p 0))
(assert (forall ((?n Int)) (=> (<= 0 ?n) (=> (p ?n) (p (+ ?n 1))))))
(assert (not (forall ((?n Int)) (=> (<= 0 ?n) (p ?n)))))
```

To translate this into a set theoretical statement, we must give a specific set representing integers. For natural numbers a reasonable option is to take the finite ordinals (the members of ω). As part of a formalization of Conway's surreal numbers [12] we also have a unary minus operation on all surreal numbers (including ordinals). The details are not important here, but it is sufficient to note that $-0 = 0$, $-n \notin \omega$ if $n \in \omega$ and $--x = x$ for all surreal numbers x. We take int to be the set $\omega \cup \{-n | n \in \omega\}$ and use int as the fixed interpretation of the sort Int. In the Megalodon preamble file we use this definition appears as follows:

```
Definition int : set := omega :\/: {- n|n :e omega}.
```

We also have a binary operation $+$ on surreal numbers which behaves as expected on int, as well as orderings $<$ and $\leq$ on surreal numbers. In general we will not give details about definitions unless they are relevant. We will only state some relevant properties we use, but emphasize that all properties we use have been previously proven in Megalodon and published

[3] http://grid01.ciirc.cvut.cz/~chad/megalodon-1.8.tgz.
[4] The full data is available at http://grid01.ciirc.cvut.cz/~chad/smt2022data.tgz.

into the Proofgold chain. There are no goals left open. To make the translation more direct on propositions, we assume $\mathcal{T}^p(s < t)$ is $\mathcal{T}(s) < \mathcal{T}(t)$ and $\mathcal{T}^p(s \leq t)$ is $\mathcal{T}(s) \leq \mathcal{T}(t)$ when s and t are of type Int.

We have chosen to locally define bp as follows:

```
Let bp : set -> prop := fun b => 0 :e b.
```

We briefly consider the behavior of bp when applied to booleans (members of the set $\{0, 1\}$). The negation of bp 0 is $0 \notin 0$ which is provable, so bp 0 acts as the false proposition. On the other hand bp 1 is $0 \in 1$ which is provable, so bp 0 acts as the true proposition. Such local definitions act more as notation that is translated away. Other definitions would also work.

The statement of the set theoretic translation of the SMT2 problem appears as follows in Megalodon:

```
Theorem example1ind_unsat:
  forall p :e 2 :^: int,
       bp (p 0)
    -> (forall n :e int, 0 <= n -> bp (p n) -> bp (p (n + 1)))
    -> ~(forall n :e int, 0 <= n -> bp (p n))
    -> False.
```

The set 2 :^: int denotes the set of functions from integers to booleans: 2^{int}. Essentially the statement says the three (translated) assertions lead to a contradiction. Note that since p 0 is a boolean (a set which is a member of $\{0, 1\}$), the coercion bp is used to create the corresponding proposition whenever necessary.

The proof in Megalodon proceeds as follows: we assume p is in the set 2^{int} and assume the three properties hold. In the preamble there is a predicate nat_p that holds for the finite ordinals. A previously proven induction principle is included:

```
nat_ind : forall p:set->prop,
     p 0
  -> (forall n, nat_p n -> p n -> p (ordsucc n))
  -> forall n, nat_p n -> p n.
```

It is straightforward to prove the translated statement from this already known induction principle.

5.2 Pigeonhole

Our second example will be two versions of the Pigeonhole Principle. We use arrays from integers to integers (with some constraints) to play the role of functions from finite ordinals to finite ordinals. In the first version we will state that every array acting as a function from $\{0, \ldots, n\}$ to $\{0, \ldots, n-1\}$ is not injective. In SMT2 format we assert the negation of this statement as follows:

```
(assert
 (not
  (forall
   ((?n Int))
   (=> (>= ?n 0)
       (forall
        ((?f (Array Int Int)))
        (=> (forall ((?i Int))
              (=> (and (<= 0 ?i) (<= ?i ?n))
                  (and (<= 0 (select ?f ?i)) (< (select ?f ?i) ?n))))
```

```
            (exists ((?i Int) (?j Int))
                (and (<= 0 ?i) (< ?i ?j) (<= ?j ?n)
                     (= (select ?f ?i) (select ?f ?j)))))))))))
```

In order to translate this SMT2 problem into a statement of formal set theory we must interpret arrays. We will translate to a statement that universally quantifies over appropriate interpretations of arrays. An interpretation of arrays is a (meta-)function *Array* taking two sets and returning a set satisfying the following property:

```
Definition Array_interp : (set -> set -> set) -> prop
  := fun Array =>
         (forall X Y, Array X Y c= Y :^: X)
      /\ (forall X Y, forall f :e Array X Y, forall x :e X, forall y :e Y,
                (fun u :e X => if u = x then y else f u) :e Array X Y).
```

That is: for sets X and Y, $Array\ X\ Y$ must be a set of functions from X to Y that is closed under changing one value.[5] To deal with arrays, we modify the translation so that $\mathcal{T}_\varphi Array$ is a special selected variable $Array : \iota\iota$ and produce the set theoretic problem

$$\forall Array.\texttt{Array\_interp}\ Array \to \forall \sigma_1 \cdots \sigma_n \in U.$$
$$\forall u_1 \in \mathcal{T}_{\varphi_1}(\alpha_1) \cdots \forall u_m \in \mathcal{T}_{\varphi_1}(\alpha_m).\mathcal{T}^p_{\varphi_2}(P_1) \to \cdots \to \mathcal{T}^p_{\varphi_2}(P_k) \to \bot.$$

Translating the Pigeonhole SMT problem to the formal set theory of Megalodon we have the following theorem:

```
Theorem PigeonHoleArrays_1_unsat :
  forall Array:set -> set -> set,
    Array_interp Array ->
    ~(forall n :e int, 0 <= n ->
      forall f :e Array int int,
         (forall i :e int, 0 <= i /\ i <= n ->
                           0 <= f i /\ f i < n)
      -> (exists i j :e int, 0 <= i /\ i < j /\ j <= n /\ f i = f j))
    -> False.
```

We can prove the set theoretic version by reducing to the following previously proven version of the Pigeonhole principle:

```
PigeonHole_nat :
 forall n, nat_p n ->
 forall f:set -> set,
       (forall i :e ordsucc n, f i :e n)
   -> ~(forall i j :e ordsucc n, f i = f j -> i = j).
```

A second version of the Pigeonhole principle states that every (array acting as an) injective function from $\{0, \ldots, n-1\}$ into $\{0, \ldots, n-1\}$ is surjective. As an SMT2 problem this can be stated as follows:

```
(assert
 (not
  (forall
   ((?n Int))
   (=> (>= ?n 0)
       (forall
```

[5] Note that this allows the set of arrays to be empty. If all types in SMT3 will be assumed to be nonempty, then this definition should be changed.

```
        ((?f (Array Int Int)))
        (=> (forall ((?i Int))
               (=> (and (<= 0 ?i) (< ?i ?n))
                   (and (<= 0 (select ?f ?i)) (< (select ?f ?i) ?n))))
            (=> (forall ((?i Int) (?j Int))
                  (=> (and (<= 0 ?i) (< ?i ?n) (<= 0 ?j) (< ?j ?n)
                           (= (select ?f ?i) (select ?f ?j)))
                      (= ?i ?j)))
                (forall
                 ((?j Int))
                 (=> (and (<= 0 ?j) (< ?j ?n))
                     (exists ((?i Int))
                       (and (<= 0 ?i) (< ?i ?n) (= (select ?f ?i) ?j))))))))))))
```

The corresponding Megalodon theorem looks as follows:

```
Theorem PigeonHoleArrays_2_unsat :
  forall Array:set -> set -> set,
    Array_interp Array ->
    ~(forall n :e int, 0 <= n ->
      forall f :e Array int int,
         (forall i :e int, 0 <= i /\ i < n ->
                           0 <= f i /\ f i < n)
      -> (forall i j :e int,
               0 <= i /\ i < n /\ 0 <= j /\ j < n /\ f i = f j
                   -> i = j)
      -> (forall j :e int, 0 <= j /\ j < n
            -> exists i :e int, 0 <= i /\ i < n /\ f i = j))
    -> False.
```

The Megalodon proof proceeds by reducing to a similar previously proven version of the Pigeonhole Principle. However, it would also be possible to infer the second version from the first version simply by instantiating with an array with one element changed.

5.3 Failure of Schroeder-Bernstein for Arrays

As a third example, we consider the Schroeder-Bernstein property for arrays. That is, we consider whether or not two types α and β must have a bijection between them if there are injections from α into β and β into α. In this case the negation of the property is satisfiable and we give an interpretation of arrays for which the property fails. Usually in logic there is either a proof on the one hand or a model on the other. However, in this case we can also give a proof term for a proof of the negation of the set theoretical property (where the negation is before the quantifier over possible interpretations of arrays).

For the SMT2 problem we let f and g be of appropriate array types and assume f and g are injective. We then assume there does not exist a bijective array.

```
(declare-fun f () (Array Int Int))
(declare-fun g () (Array Int Int))
(assert (forall ((?m Int) (?n Int)) (=> (= (select f ?m) (select f ?n)) (= ?m ?n))))
(assert (forall ((?m Int) (?n Int)) (=> (= (select g ?m) (select g ?n)) (= ?m ?n))))
(assert
  (not (exists ((?h (Array Int Int)))
         (and (forall ((?m Int) (?n Int)) (=> (= (select ?h ?m) (select ?h ?n)) (= ?m ?n)))
              (forall ((?n Int)) (exists ((?m Int)) (= (select ?h ?m) ?n)))))))
```

The translation of this problem to a set theoretic proposition in Megalodon appears as follows:

```
  forall Array:set -> set -> set,
     Array_interp Array ->
     (forall f g :e Array int int,
         (forall m n :e int, f m = f n -> m = n)
      -> (forall m n :e int, g m = g n -> m = n)
      -> ~(exists h :e Array int int,
               (forall m n :e int, h m = h n -> m = n)
            /\ (forall n :e int, exists m :e int, h m = n))
      -> False)).
```

This is not provable. However, we can prove the negation of the proposition (if we are careful to put the negation before the quantifier for the interpretation of arrays).

```
Theorem SchroederBernsteinArrays_sat :
  ~(forall Array:set -> set -> set,
     Array_interp Array ->
     (forall f g :e Array int int,
         (forall m n :e int, f m = f n -> m = n)
      -> (forall m n :e int, g m = g n -> m = n)
      -> ~(exists h :e Array int int,
               (forall m n :e int, h m = h n -> m = n)
            /\ (forall n :e int, exists m :e int, h m = n))
      -> False)).
```

The most important choice for proving this negated proposition is properly instantiating for *Array*. We start by defining an injective function from integers to natural numbers which sends negative integers x to $(2(-x)) + 1$ and nonnegative integers x to $2x$.

```
set int_into_nat : set := (fun x :e int => if x < 0 then ordsucc (2 * (- x)) else 2 * x).
```

We can now inductively define the collection of all functions that are the same as `int_into_nat` except on finitely many elements.

```
set ArrayIntInt_p : set -> prop := fun f =>
  forall p:set -> prop,
        p int_into_nat
     -> (forall f, forall x y :e int, p f -> p (fun u :e int => if u = x then y else f u))
     -> p f.
```

Finally we can define *Array* (the term we will use as the instantiation for the quantified variable *Array*) to be the set of all functions unless both arguments are the set of integers, in which case the functions must satisfy `ArrayIntInt_p`.

```
set Array : set -> set -> set :=
  fun A B =>
     if A = int /\ B = int then
       {f :e int :^: int | ArrayIntInt_p f}
     else B :^: A.
```

Intuitively it should be clear that this choice satisfies `Array_interp`. It is also the case that *Array* int int contains no bijection. Formally we prove that every function satisfying `ArrayIntInt_p` has a lower bound and then use this to conclude that such a function cannot be a surjection.

5.4 Integer Difference Logic

Our final examples will be two integer difference logic problems from the "job shop" collection from QF_IDL portion of the SMT library. One is satisfiable and the other is unsatisfiable. In both cases we can obtain proof terms for the corresponding set theoretic proposition.

As described in [18] satisfiability of a set of atoms of the form $x_1 + -x_0 \leq v_0$, $x_2 + -x_1 \leq v_1$, $\ldots$, $x_0 + -x_{n-1} \leq v_{n-1}$ (where the variables range over integers) can be decided by forming a certain directed graph with edges labeled by integers and checking if there is a negative loop. If there is no negative loop, then values for the variables can be computed from the graph.

We first consider the problem `jobshop2-2-1-1-4-4-11` (slightly modified to be more readable). In the problem there are five integer variables s_1^1, s_2^1, s_1^2, s_2^2 and ref. The assertion given in the problem

$$(v_0 \vee v_1) \wedge (v_2 \vee v_3) \wedge s_2^1 - s_1^1 \geq 4 \wedge s_2^2 - s_1^2 \geq 4$$
$$\wedge s_2^1 - \text{ref} \leq 7 \wedge s_2^2 - \text{ref} \leq 7 \wedge s_1^1 - \text{ref} \geq 0 \wedge s_1^2 - \text{ref} \geq 0$$

where v_1, v_0, v_3 and v_2 are locally defined (via a `let`) to be the atoms $s_1^1 - s_1^2 \geq 4$, $s_1^2 - s_1^1 \geq 4$, $s_2^1 - s_2^2 \geq 4$ and $s_2^2 - s_2^1 \geq 4$, respectively. This problem is unsatisfiable. An informal proof of unsatisfiability proceeds by splitting into two cases via the disjunction $v_3 \vee v_2$. In the v_3 case there is a negative loop given by $s_2^1, s_2^2, s_1^2, \text{ref}$. In the v_2 case there is a negative loop given by $s_2^2, s_2^1, s_1^1, \text{ref}$.

The set theoretic version of the problem can be defined as the following proposition in Megalodon.

```
Definition jobshop2_2_1_1_4_4_11 : prop :=
    forall s1_1 s1_2 s2_1 s2_2 ref :e int,
    forall v_0:prop, v_0 = (4 <= s2_1 + - s1_1)
 -> forall v_1:prop, v_1 = (4 <= s1_1 + - s2_1)
 -> forall v_2:prop, v_2 = (4 <= s2_2 + - s1_2)
 -> forall v_3:prop, v_3 = (4 <= s1_2 + - s2_2)
 -> ((v_0 \/ v_1) /\ (v_2 \/ v_3)
  /\ 4 <= s1_2 + - s1_1 /\ 4 <= s2_2 + - s2_1
  /\ s1_2 + - ref <= 7 /\ s2_2 + - ref <= 7
  /\ 0 <= s1_1 + - ref /\ 0 <= s2_1 + - ref)
 -> False.
```

Note that we have slightly modified the inequalities to all use $\leq$ for simplicity. Also, we combine the unary $-$ with the binary $+$ operator instead of using a binary $-$ operator. Finally, we have replaced the `let` declarations for the v_i's with universally quantified variables and an assumed identity. The proposition above corresponds to the unsatisfiability of the original SMT problem and it can be proven in set theory using the negative loops mentioned above and the following (formally proven) result.

```
Theorem idl_negcycle_4 : forall x y z w v1 v2 v3 v4,
    SNo x -> SNo y -> SNo z -> SNo w
 -> SNo v1 -> SNo v2 -> SNo v3 -> SNo v4
 -> v1 + v2 + v3 + v4 < 0
 -> y + - x <= v1 -> z + - y <= v2
 -> w + - z <= v3 -> x + - w <= v4
 -> False.
```

The theorem `idl_negcycle_4` is specific to negative loops of length 4, but is also more general since variables range over values satisfying the predicate `SNo`, a predicate true for integers, real

numbers, and more (Conway's extension of the real numbers described in [12]). While it is relatively easy to prove `idl_negcycle_4` directly, it is somewhat unsatisfying to have the result be specific to cycles of length 4. Fortunately the theorem can be easily proven as a consequence of the following (formally proven) general result about cycles of length $n+1$, by induction on n.

```
Theorem SNo_idl_cycle_nonneg : forall n, nat_p n ->
  forall f g:set -> set,
             (forall i :e ordsucc n, SNo (f i))
           -> (forall i :e ordsucc n, SNo (g i))
           -> f (ordsucc n) = f 0
           -> (forall i :e ordsucc n, f (ordsucc i) + - f i <= g i)
           -> 0 <= finite_add_SNo (ordsucc n) g.
```

Note that f corresponds to an $n+2$-tuple (with $f(n+1) = f0$) and g corresponds to an $n+1$-tuple. The function `finite_add_SNo` takes a natural number n and a function g and returns the sum $\Sigma_{i \in n} gi$. To infer the special case for 4-cycles we apply the general result with 3 for n, an f given by the 5-tuple (x, y, z, w, x) and an g given by the 4-tuple $(v1, v2, v3, v4)$.

The final example we consider is `jobshop2-2-1-1-4-4-12` (again slightly modified) which is a simple modification of the previous example by changing each 7 to 8.

$$(v_0 \vee v_1) \wedge (v_2 \vee v_3) \wedge s_2^1 - s_1^1 \geq 4 \wedge s_2^2 - s_1^2 \geq 4$$
$$\wedge s_2^1 - \text{ref} \leq 8 \wedge s_2^2 - \text{ref} \leq 8 \wedge s_1^1 - \text{ref} \geq 0 \wedge s_1^2 - \text{ref} \geq 0$$

This makes the problem satisfiable by taking $s_1^1 = 0$, $s_2^1 = 4$, $s_1^2 = 4$, $s_2^2 = 8$ and $\text{ref} = 0$.

The corresponding set theoretic proposition is given as follows:

```
Definition jobshop2_2_1_1_4_4_12 : prop :=
    forall s1_1 s1_2 s2_1 s2_2 ref :e int,
    forall v_0:prop, v_0 = (4 <= s2_1 + - s1_1)
 -> forall v_1:prop, v_1 = (4 <= s1_1 + - s2_1)
 -> forall v_2:prop, v_2 = (4 <= s2_2 + - s1_2)
 -> forall v_3:prop, v_3 = (4 <= s1_2 + - s2_2)
 -> ((v_0 \/ v_1) /\ (v_2 \/ v_3)
  /\ 4 <= s1_2 + - s1_1 /\ 4 <= s2_2 + - s2_1
  /\ s1_2 + - ref <= 8 /\ s2_2 + - ref <= 8
  /\ 0 <= s1_1 + - ref /\ 0 <= s2_1 + - ref)
 -> False.
```

Since the problem is satisfiable, the proposition cannot be proven. However, we can prove its negation by assuming the proposition holds and applying it to the values above. One must then proven v_0 and v_2 hold (giving all the disjunctions) and prove the remaining inequalities hold.

6 Conclusion

A finitely axiomatized set theory using Church's type theory instead of first-order logic provides a simple way of obtaining a candidate semantics for SMT3. Via Curry-Howard, we also naturally obtain a candidate notion of independently checkable proof term.

Acknowledgements

The results were supported by the Ministry of Education, Youth and Sports within the dedicated program ERC CZ under the project POSTMAN no. LL1902. This scientific article is part of the RICAIP project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857306.

References

[1] SMT-LIB version 3.0 - preliminary proposal, 2021. http://smtlib.cs.uiowa.edu/version3.shtml.

[2] Peter Aczel. On relating type theories and set theories. In Thorsten Altenkirch, Wolfgang Naraschewski, and Bernhard Reus, editors, *TYPES*, volume 1657 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 1998.

[3] P. B. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof*. Kluwer Academic Publishers, 2nd edition, 2002.

[4] Peter B. Andrews. General models and extensionality. *J. Symb. Log.*, 37:395–397, 1972.

[5] Bruno Barras. Sets in Coq, Coq in Sets. *Journal of Formalized Reasoning*, 3(1), 2010.

[6] Clark Barrett, Pascal Fontaine, and Cesare Tinelli. The SMT-LIB Standard: Version 2.6. Technical report, Department of Computer Science, The University of Iowa, 2017. Available at `www.SMT-LIB.org`.

[7] Clark Barrett, Pascal Fontaine, and Cesare Tinelli. New working group on SMT proofs, 2021. https://groups.google.com/g/smt-lib/c/a0-U-nU4J68.

[8] Yves Bertot and Pierre Castéran. *Interactive Theorem Proving and Program Development. Coq'Art: The Calculus of Inductive Constructions*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.

[9] Chad E. Brown. Notes on semantics of and proofs for smt, Sep 2021.

[10] Chad E. Brown and Karol Pąk. A tale of two set theories. In Cezary Kaliszyk, Edwin C. Brady, Andrea Kohlhase, and Claudio Sacerdoti Coen, editors, *Intelligent Computer Mathematics - 12th International Conference, CICM 2019, Prague, Czech Republic, July 8-12, 2019, Proceedings*, volume 11617 of *Lecture Notes in Computer Science*, pages 44–60. Springer, 2019.

[11] Alonzo Church. A formulation of the simple theory of types. *The Journal of Symbolic Logic*, 5:56–68, 1940.

[12] John H. Conway. *On numbers and games, Second Edition*. A K Peters, 2001.

[13] N. de Bruijn. The Mathematical language AUTOMATH, its usage, and some of its extensions. In *Symposium on Automatic Demonstration*, pages 29–61. Lecture Notes in Mathematics, 125, Springer, 1970.

[14] N .G. de Bruijn. A survey of the project AUTOMATH. In J. P. Seldin and J. R. Hindley, editors, *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, pages 579–606. Academic Press, 1980.

[15] Leon Henkin. Completeness in the theory of types. *The Journal of Symbolic Logic*, 15:81–91, 1950.

[16] W.A. Howard. The formulas-as-types notion of construction. In J.P. Seldin and J.R. Hindley, editors, *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 479–490, New York, 1980. Academic Press.

[17] B. Jacobs. *Categorical Logic and Type Theory*. ISSN. Elsevier Science, 1999.

[18] Hyondeuk Kim and Fabio Somenzi. Finite instantiations for integer difference logic. In *2006 Formal Methods in Computer Aided Design*, pages 31–38. IEEE, nov 2006.

[19] Dominik Kirst and Gert Smolka. Categoricity results and large model constructions for second-order ZF in dependent type theory. *Journal of Automated Reasoning*, 2018. First Online: 11 October 2018.

[20] Joachim Lambek and Philip Scott. *Introduction to higher order categorical logic*. Cambridge University Press, Cambridge, UK, 1986.

[21] Gyesik Lee and Benjamin Werner. Proof-irrelevant model of CC with predicative induction and judgmental equality. *Logical Methods in Computer Science*, 7(4), 2011.

[22] Zhaohui Luo. *Computation and reasoning: a type theory for computer science*. Oxford University Press, Inc., New York, NY, USA, 1994.

[23] The Coq development team. *The Coq proof assistant reference manual*. LogiCal Project, 2020. Version 8.12.

[24] Dag Prawitz. *Natural deduction: a proof-theoretical study*. Dover, 2006.

[25] M.H.B. Sørensen and P. Urzyczyn. *Lectures on the Curry-Howard Isomorphism*. Rapport (Københavns universitet. Datalogisk institut). Datalogisk Institut, Københavns Universitet, 1998.

[26] A. Tarski. Der Aussagenkalkül und die Topologie. *Fundamentae Mathematicae*, 31:103–134, 1938.

[27] Benjamin Werner. Sets in types, types in sets. In Martín Abadi and Takayasu Ito, editors, *TACS*, volume 1281 of *Lecture Notes in Computer Science*, pages 530–346. Springer, 1997.

A appendix - to do