Formalizing Mathematical Knowledge as a Biform Theory Graph: A Case Study*

Jacques Carette and William M. Farmer

Computing and Software, McMaster University, Canada
http://www.cas.mcmaster.ca/~carette
http://imps.mcmaster.ca/wmfarmer

Abstract. A *biform theory* is a combination of an axiomatic theory and an algorithmic theory that supports the integration of reasoning and computation. These are ideal for formalizing algorithms that manipulate mathematical expressions. A *theory graph* is a network of *theories* connected by meaning-preserving *theory morphisms* that map the formulas of one theory to the formulas of another theory. Theory graphs are in turn well suited for formalizing mathematical knowledge at the most convenient level of abstraction using the most convenient vocabulary. We are interested in the problem of whether a body of mathematical knowledge can be effectively formalized as a theory graph of biform theories. As a test case, we look at the graph of theories encoding natural number arithmetic. We used two different formalisms to do this, which we describe and compare. The first is realized in CTT_{uqe}, a version of Church's type theory with quotation and evaluation, and the second is realized in Agda, a dependently typed programming language.

1 Introduction

There are many methods for encoding mathematical knowledge. The two most prevalent are the *axiomatic* and the *algorithmic*. The axiomatic method, famously employed by Euclid in his *Elements* circa 300 BCE, encodes a body of knowledge as an *axiomatic theory* composed of a language and a set of *axioms* expressed in that language. The axioms are assumptions about the *concepts* of the language and the logical consequences of the axioms are the *facts* about the concepts. The algorithmic method in contrast uses an *algorithmic theory*, composed of a language and a set of *algorithms* that perform symbolic computations over the expressions of the language. Each algorithm procedurally encodes its input/output relation. For example, an algorithm that symbolically adds expressions that represent rational numbers encodes the addition function $+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ over the rational numbers.

A complex body of mathematical knowledge comprises many different theories; these can be captured by the *little theories method* [10] as a *theory graph* [13] consisting of theories as nodes and theory morphisms as directed edges. A theory morphism is a meaning-preserving mapping from the formulas of one theory

* This research was supported by NSERC.

to the formulas of another. The theories serve as abstract mathematical models and the morphisms serve as information conduits that enable definitions and theorems to be transported from one theory to another [2]. A theory graph enables mathematical knowledge to be formalized at the most convenient level of abstraction using the most convenient vocabulary. Moreover, the structure of a theory graph provides the means to access relevant concepts and facts (c&f), reduce the duplication of c&f, and enable c&f to be interpreted in multiple ways.

The axiomatic method is the basis for formalizing mathematical knowledge in proof assistants and logical frameworks. Although many proof assistants support the little theories method to some extent, very few provide the means to explicitly build theory graphs. Notable exceptions are the IMPS theorem proving system [11] and the MMT module system for mathematical theories [16].

Computer algebra systems on the other hand are based on algorithmic theories, which are not usually organized as a graph. An exception is the Axiom system [12] in which a network of abstract and concrete algorithmic theories are represented by Axiom categories and domains, respectively. Algorithmic theories are challenging to fully formalize because a specification of a symbolic algorithm that encodes a mathematical function requires the ability to talk about the relationship between syntax and semantics.

Axiomatic and algorithmic knowledge complement each other, and both are needed. A *biform theory* [3, 5] combines both, and furthermore supports the integration of reasoning and computation. We argue in [3] that biform theories are needed to build *high-level theories* analogous to high-level programming languages. Biform theories are challenging to formalize for the same reasons that algorithmic theories are challenging to formalize.

We are interested in the problem of whether the little theories method can be applied to biform theories. That is, can a body of mathematical knowledge be effectively formalized as a theory graph of biform theories? We use a graph (of biform theories) encoding natural number arithmetic as a test case. We describe two different formalizations, and compare the results. The first formalization is realized using the global approach in CTT_{uqe} [9], a variant of CTT_{qe} [7, 8], a version of Church's type theory with quotation and evaluation, while the second is realized using the local approach in Agda [14, 15] , a dependently typed programming language. This dual formalization, contrasting the two approaches, forms the core of our contribution; each formalization has some smaller contributions, some of which may be of independent interest.

The rest of the paper is organized as follows. The notion of a biform theory is defined and discussed in section 2. The theories that encode natural number arithmetic are presented in section 3. The CTT_{uqe} formalization is discussed in section 4, and the Agda version in section 5. These two are presented in full in appendices A and B of [4]. Section 6 compares the two formalizations. The paper ends with conclusions and future work in section 7.

The authors are grateful to the reviewers for their comments and suggestions.

2 Biform Theories

Let $\mathcal{E}$ be a set of expressions and $f : \mathcal{E}^n \to \mathcal{E}$ be an n-ary function where $n \geq 1$. A *transformer for* f is an algorithm that implements f. Transformers manipulate expressions e in various ways: simple ones build bigger expressions from pieces, select components of e, or check whether e satisfies some syntactic property. More sophisticated transformers manipulate expressions in a mathematically meaningful way. We call these kinds of transformers *syntax-based mathematical algorithms (SBMAs)* [6]. Examples include algorithms that apply arithmetic operations to numerals, factor polynomials, transpose matrices, and symbolically differentiate expressions with variables. The *computational behavior* of a transformer is the relationship between its input and output expressions. When the transformer is an SBMA, its *mathematical meaning* is the relationship between the mathematical meanings of its input and output expressions.

A *biform theory* T is a triple (L, Π, Γ) where L is a language of some underlying logic, Π is a set of transformers for functions over expressions of L, and Γ is a set of formulas of L. L is generated from a set of symbols that include, e.g., types and constants. Each symbol is the name for a concept of T. The transformers in Π are for functions represented by symbols of L. The members of Γ are the *axioms* of T. They specify the concepts of T including the computational behaviors of transformers and the mathematical meanings of SBMAs. The underlying logic provides the semantic foundation for T. We say T is an *axiomatic theory* if Π is empty and an *algorithmic theory* if Γ is empty.

Expressing a biform theory in the underlying logic requires infrastructure for reasoning about expressions manipulated by the transformers as syntactic entities. The infrastructure provides a basis for *metareasoning with reflection* [7]. There are two main approaches for obtaining this infrastructure [6]. The *local approach* is to produce a deep embedding of a sublanguage L' of L that include all the expressions manipulated by the transformers of Π. The deep embedding consists of (1) an inductive type of *syntactic values* that represent the syntactic structures of the expressions in L', (2) an *informal quotation operator* that maps the expressions in L' to syntactic values, and (3) a *formal evaluation operator* that maps syntactic values to the values of the expressions in L' that they represent. The *global approach* is to replace the underlying logic of L with a logic such as that of [7] that has (1) an inductive type of *syntactic values* for all the expressions in L, (2) a *global formal quotation operator*, and (3) a *global formal evaluation operator*.

There are several ways, in a proof assistant, to construct a transformer π for $f : \mathcal{E}^n \to \mathcal{E}$. The simplest is to define f as a lambda abstraction A_f, and then π computes the value $f(e_1, \ldots, e_n)$ by reducing $A_f(e_1, \ldots, e_n)$ using β-reduction (and possibly other transformations such as δ-reduction, etc). Another method is to specify the computational behavior of f by axioms, and then π can be implemented as a tactic that applies the axioms to $f(e_1, \ldots, e_n)$ as, e.g., rewrite rules or conditional rewrite rules. Finally, the computational behavior or mathematical meaning of f can be specified by axioms, and then π can be a

program which satisfies these axioms; this program can operate on either internal or external data structures representing the expressions $e_1, \ldots, e_n$.

3 Natural Number Arithmetic: A Test Case

Figure 1 shows a theory graph composed of biform theories encoding natural number arithmetic. We start with eight axiomatic theories (seven in first-order logic (FOL) and one in simple type theory (STT)) and then add a variety of useful transformers in the appropriate theories. These eight are chosen because they fit together closely and have simple axiomatizations. Of the first-order theories, BT1 and BT5 are theories of 0 and S (which denotes the successor function); BT2 and BT6 are theories of 0, S, and $+$; and BT3, BT4, and BT7 are theories of 0, S, $+$, and $*$. Several other biform theories could be added to this graph, most notably Skolem arithmetic, the complete theory of 0, S, and $*$, which has a very complicated axiomatization [17]. The details of each theory is given below.

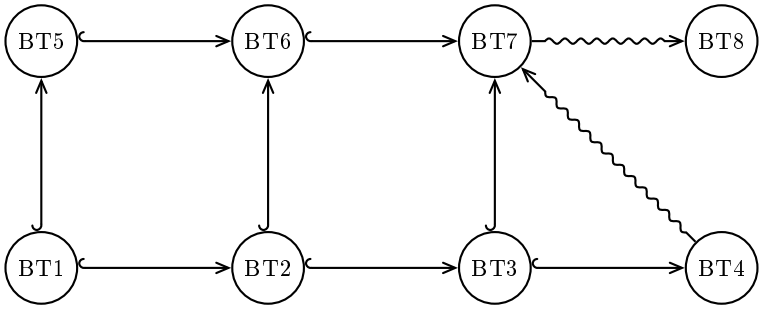

Fig. 1. Biform Theory Graph Test Case

Figure 1 shows the morphisms that connect these theories. The $\hookrightarrow$ arrows denote strict theory inclusions. The morphism from BT4 to BT7 is the identity mapping. It is meaning-preserving since each axiom of BT4 is a theorem of BT7. In particular, A7 follows from the induction schema A10. The theory morphism from BT7 to BT8 is interlogical since their logics are different. It is defined by the mapping of $0, S, +, *$ to $0_\iota, S_{\iota\to\iota}, +_{\iota\to\iota\to\iota}, *_{\iota\to\iota\to\iota}$, respectively, where $+_{\iota\to\iota\to\iota}$ and $*_{\iota\to\iota\to\iota}$ are defined constants in BT8. It is meaning-preserving since A1–A6 and the instances of the induction schema A10 map to theorems of BT8.

We have formalized this biform theory graph in two ways: the first in CTT_{uqe} using the global approach and the second in Agda using the local approach. These are discussed in the next two sections, while the full details are given in appendices A and B of [4]. A "conventional" mathematical presentation of the theories would be as follows.

Biform Theory 1 (BT1: Simple Theory of 0 and S)
Logic: FOL. *Constants*: 0 (0-ary), S (unary).

Axioms:
A1. $S(x) \neq 0$.
A2. $S(x) = S(y) \supset x = y$.
Properties: incomplete, undecidable.
Transformers: Recognizer for the formulas of the theory.

Biform Theory 2 (BT2: Simple Theory of 0, S, and $+$)
Extends BT1.
Logic: FOL. *Constants*: $+$ (binary, infix).
Axioms:
A3. $x + 0 = x$.
A4. $x + S(y) = S(x + y)$.
Properties: incomplete, undecidable.
Transformers: Recognizer for the formulas of the theory and algorithm for adding natural numbers as binary numerals.

Biform Theory 3 (BT3: Simple Theory of 0, S, $+$, and $*$)
Extends BT2.
Logic: FOL. *Constants*: $*$ (binary, infix).
Axioms:
A5. $x * 0 = 0$.
A6. $x * S(y) = (x * y) + x$.
Properties: incomplete, undecidable.
Transformers: Recognizer for the formulas of the theory and algorithm for multiplying natural numbers as binary numerals.

Biform Theory 4 (BT4: Robinson Arithmetic (Q))
Extends BT3.
Logic: FOL.
Axioms:
A7. $x = 0 \vee \exists\, y\ .\ S(y) = x$.
Properties: essentially incomplete, essentially undecidable.

Biform Theory 5 (BT5: Complete Theory of 0 and S)
Extends BT1.
Logic: FOL.
Axioms:
A8. $(A(0) \wedge \forall\, x\ .\ (A(x) \supset A(S(x)))) \supset \forall\, x\ .\ A(x)$
where A is any formula of BT5 in which x is not bound and $A(t)$ is the result of replacing each free occurrence of x in A with the term t.
Properties: complete, decidable.
Transformers: Generator for instances of the theory's induction schema and decision procedure for the theory.

Biform Theory 6 (BT6: Presburger Arithmetic)
Extends BT2 and BT5.
Logic: FOL.

Axioms:

A9. $(A(0) \wedge \forall\, x\, .\, (A(x) \supset A(S(x)))) \supset \forall\, x\, .\, A(x)$

where A is any formula of BT6 in which x is not bound and $A(t)$ is the result of replacing each free occurrence of x in A with the term t.

Properties: complete, decidable.

Transformers: Generator for instances of the theory's induction schema and decision procedure for the theory.

Biform Theory 7 (BT7: First-Order Peano Arithmetic)
Extends BT3 and BT6.

Logic: FOL.

Axioms:

A10. $(A(0) \wedge \forall\, x\, .\, (A(x) \supset A(S(x)))) \supset \forall\, x\, .\, A(x)$

where A is any formula of BT7 in which x is not bound and $A(t)$ is the result of replacing each free occurrence of x in A with the term t.

Properties: essentially incomplete, essentially decidable.

Transformers: Generator for instances of the theory's induction schema.

Biform Theory 8 (BT8: Higher-Order Peano Arithmetic)
Logic: STT. *Types*: ι. *Constants*: 0_ι, $S_{\iota\to\iota}$.

Axioms:

A11. $S_{\iota\to\iota}(x_\iota) \neq 0_\iota$.

A12. $S_{\iota\to\iota}(x_\iota) = S_{\iota\to\iota}(y_\iota) \supset x_\iota = y_\iota$.

A13. $(p_{\iota\to o}(0) \wedge \forall\, x_\iota\, .\, (p_{\iota\to o}(x_\iota) \supset p_{\iota\to o}(S(x_\iota)))) \supset \forall\, x_\iota\, .\, p_{\iota\to o}(x_\iota)$.

Properties: essentially incomplete, essentially decidable, categorical for standard models.

It is important to note that axioms A8, A9 and A10 are all different since they are over different languages; in particular, **BT6** adds $+$ to the language of **BT5**, and **BT7** adds $*$ to the language of **BT6**.

4 Study 1: Test Case Formalized in CTT_{uqe}

CTT_{uqe} supports the global approach for metareasoning with reflection. CTT_{uqe} contains (1) a logical base type ϵ that is an inductive type of syntactic values called *constructions* which are expressions of type ϵ, (2) a global quotation operator $\ulcorner\cdot\urcorner$ that maps each expression $\mathbf{A}_\alpha$ of CTT_{uqe} to a construction that represents the syntactic structure of $\mathbf{A}_\alpha$, and (3) a typed global evaluation operator $[\![\cdot]\!]_\alpha$ that maps each construction $\mathbf{B}_\epsilon$ of CTT_{uqe} representing an expression $\mathbf{A}_\alpha$ of type α to an expression whose value is the same as $\mathbf{A}_\alpha$. See [9] for details.

A *biform theory* of CTT_{uqe} is a triple (L, Π, Γ) where L is a language generated by a set of base types and constants of CTT_{uqe}, Π is a set of transformers over expressions of L, and Γ is a set of formulas of L. Each transformer is for a constant in L whose type has the form $\epsilon \to \cdots \to \epsilon$. We present biform theories in CTT_{uqe} as a set of base types, constants, axioms, transformers, and theorems. The base types are divided into primitive and defined base types. A defined base

type is declared by a formula that equates the base type to a nonempty subset of some type of L. Similarly, the constants are divided into primitive and defined constants. A defined constant $\mathbf{c}_\alpha$ is declared by an equation $\mathbf{c}_\alpha = \mathbf{A}_\alpha$ where $\mathbf{A}_\alpha$ is a defined expression.

The biform theory graph test case given in section 3 is formalized in CTT_{uqe} as a theory graph of eight CTT_{uqe} theories as shown in appendix A of [4]. Since CTT_{uqe} is not currently implemented, it is not possible to give the transformers as implemented algorithms. Instead we described their intended behavior.

We will concentrate our discussion on BT6 (given below). We have not included the following components of BT6 (that should be in BT6 according to its definition in section 3) that are redundant or are subsumed by other components: Constants $\mathsf{BT5\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}$, $\mathsf{IS\text{-}FO\text{-}BT1}_{\epsilon\to\epsilon}$, and $\mathsf{IS\text{-}FO\text{-}BT1\text{-}ABS}_{\epsilon\to\epsilon}$; axioms 27 and 28; and transformers π_1, π_2, π_{11}, π_{12}, and π_{13}. See [9] for details.

BT6 has the usual constants (0_ι, $S_{\iota\to\iota}$, and $+_{\iota\to\iota\to\iota}$) and axioms (axioms 1–4 and 29) of Presburger arithmetic. Axiom 29 is the direct formalization of A9, the induction schema for Presburger arithmetic, stated in section 3. It is expressed as a single universal formula in CTT_{uqe} that ranges over constructions representing function abstractions of the form $\lambda\,\mathbf{x}_\iota\,.\,\mathbf{A}_o$. These constructions are identified by the transformers π_{15} and π_{16} for the defined constant $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$. π_{15} works by accessing data about variables, constants, and other subexpressions stored in the data structure for an expression, while π_{16} works by expanding the definition of $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$. π_{15} is sound if the definition expansion mechanism is sound. Showing the soundness of π_{14} would require a formal verification of the implementation of the data structure for expressions. Of course, the results of π_{14} could be checked using π_{16}.

This biform theory has a defined constant $\mathsf{bnat}_{\iota\to\iota\to\iota}$ with the usual base 2 notation for expressing natural numbers in a binary form. There is a constant $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$ specified by axioms 5–15 for adding quotations of these natural numbers in binary form. $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$ is implemented by transformers π_3 and π_4. π_3 is some efficient algorithm implemented outside of CTT_{uqe}, and π_4 is an algorithm that uses axioms 5–15 as conditional rewrite rules. π_4 is sound if the rewriting mechanism is sound. Showing the soundness of π_3 would require a formal verification of its program. The meaning formula for $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$, theorem 3, follows from axioms 5–15.

This biform theory also has a transformer π_{14} for $\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}$ that implements an efficient decision procedure for the first-order formulas of the theory that is specified by axiom 30. The first-order formulas of the theory are identified by the transformers π_5 and π_6 for the defined constant $\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}$ that are analogous to the transformers π_{15} and π_{16} for $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$.

Biform Theory 6 (BT6: Presburger Arithmetic)

Primitive Base Types

1. ι (type of natural numbers).

Primitive Constants

1. 0_ι.
2. $S_{\iota\to\iota}$.

3. $+_{\iota\to\iota\to\iota}$ (infix).
4. $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$ (infix).
6. $\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}$.

Defined Constants (selected)

1. $1_\iota = S\,0_\iota$.
3. $\mathsf{bnat}_{\iota\to\iota\to\iota} = \lambda\, x_\iota\,.\,\lambda\, y_\iota\,.\,((x_\iota + x_\iota) + y_\iota)$.
   Notational definition:
   $(0)_2 = \mathsf{bnat}\,0_\iota\,0_\iota$.
   $(1)_2 = \mathsf{bnat}\,0_\iota\,1_\iota$.
   $(a_1 \cdots a_n 0)_2 = \mathsf{bnat}\,(a_1 \cdots a_n)_2\,0_\iota$ where each a_i is 0 or 1.
   $(a_1 \cdots a_n 1)_2 = \mathsf{bnat}\,(a_1 \cdots a_n)_2\,1_\iota$ where each a_i is 0 or 1.
4. $\mathsf{is\text{-}bnum}_{\epsilon\to o} = \mathrm{I}\,f_{\epsilon\to o}\,.\,\forall\, u_\epsilon\,.\,(f_{\epsilon\to\epsilon}\,u_\epsilon \equiv$
   $\exists\, v_\epsilon\,.\,\exists\, w_\epsilon\,.\,(u_\epsilon = \ulcorner \mathsf{bnat}\,\lfloor v_\epsilon\rfloor\,\lfloor w_\epsilon\rfloor\urcorner \wedge$
   $(v_\epsilon = \ulcorner 0_\iota\urcorner \vee f_{\epsilon\to\epsilon}\,v_\epsilon) \wedge (w_\epsilon = \ulcorner 0_\iota\urcorner \vee w_\epsilon = \ulcorner 1_\iota\urcorner)))$.[1]
5. $\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon} = \lambda\, x_\epsilon\,.\,\mathbf{B}_\epsilon$ where $\mathbf{B}_\epsilon$ is a complex expression such that $(\lambda\, x_\epsilon\,.\,\mathbf{B}_\epsilon)\ulcorner\mathbf{A}_\alpha\urcorner$ equals $\ulcorner T_o\urcorner$ [$\ulcorner F_o\urcorner$] if $\mathbf{A}_\alpha$ is [not] a term or formula of first-order logic with equality whose variables are of type ι and whose nonlogical constants are members of $\{0_\iota, S_{\iota\to\iota}, +_{\iota\to\iota\to\iota}\}$.
7. $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon} =$
   $\lambda\, x_\epsilon\,.\,(\mathsf{if}\ (\mathsf{is\text{-}abs}_{\epsilon\to o}\,x_\epsilon)\ (\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}\,(\mathsf{abs\text{-}body}_{\epsilon\to\epsilon}\,x_\epsilon))\ \ulcorner F_o\urcorner)$.

Axioms

1. $S\,x_\iota \neq 0_\iota$.
2. $S\,x_\iota = S\,y_\iota \supset x_\iota = y_\iota$.
3. $x_\iota + 0_\iota = x_\iota$.
4. $x_\iota + S\,y_\iota = S\,(x_\iota + y_\iota)$.
5. $\mathsf{is\text{-}bnum}\,u_\epsilon \supset u_\epsilon\ \mathsf{BPLUS}\ \ulcorner(0)_2\urcorner = u_\epsilon$.

   ⋮

15. $(\mathsf{is\text{-}bnum}\,u_\epsilon \wedge \mathsf{is\text{-}bnum}\,v_\epsilon) \supset$
    $\ulcorner\mathsf{bnat}\,\lfloor u_\epsilon\rfloor\,1_\iota\urcorner\ \mathsf{BPLUS}\ \ulcorner\mathsf{bnat}\,\lfloor v_\epsilon\rfloor\,1_\iota\urcorner =$
    $\ulcorner\mathsf{bnat}\,\lfloor(u_\epsilon\ \mathsf{BPLUS}\ v_\epsilon)\ \mathsf{BPLUS}\ \ulcorner(1)_2\urcorner\rfloor\,0_\iota\urcorner$.
29. Induction Schema for S and $+$
    $\forall\, f_\epsilon\,.\,((\mathsf{is\text{-}expr}^{\iota\to o}_{\epsilon\to o}\,f_\epsilon \wedge [\![\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}\,f_\epsilon]\!]_o) \supset$
    $(([\![f_\epsilon]\!]_{\iota\to o}\,0_\iota \wedge (\forall\, x_\iota\,.\,[\![f_\epsilon]\!]_{\iota\to o}\,x_\iota \supset [\![f_\epsilon]\!]_{\iota\to o}\,(\mathsf{S}\,x_\iota))) \supset \forall\, x_\iota\,.\,[\![f_\epsilon]\!]_{\iota\to o}\,x_\iota))$.
30. Meaning formula for $\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}$.
    $\forall\, u_\epsilon\,.\,((\mathsf{is\text{-}expr}^{o}_{\epsilon\to o}\,u_\epsilon \wedge \mathsf{is\text{-}closed}_{\epsilon\to o}\,u_\epsilon \wedge [\![\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}\,u_\epsilon]\!]_o) \supset$
    $((\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}\,u_\epsilon = \ulcorner T_o\urcorner \vee \mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}\,u_\epsilon = \ulcorner F_o\urcorner) \wedge$
    $[\![\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon}\,u_\epsilon]\!]_o = [\![u_\epsilon]\!]_o))$.

Transformers

3. π_3 for $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$ is an efficient program that satisfies Axioms 5–15.
4. π_4 for $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$ uses Axioms 5–15 as conditional rewrite rules.
5. π_5 for $\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}$ is an efficient program that accesses the data stored in the data structures that represent expressions.
6. π_6 for $\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}$ uses the definition of $\mathsf{IS\text{-}FO\text{-}BT2}_{\epsilon\to\epsilon}$.

[1] Notation of the form $\ulcorner\cdots\lfloor\cdot\rfloor\cdots\urcorner$ represents a quasiquotation; see [7] for details.

14. π_{14} for $\mathsf{BT6\text{-}DEC\text{-}PROC}_{\epsilon\to\epsilon\to\epsilon}$ is an efficient decision procedure that satisfies Axiom 30.
15. π_{15} for $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$ is an efficient program that accesses the data stored in the data structures that represent expressions.
16. π_{16} for $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$ uses the definition of $\mathsf{IS\text{-}FO\text{-}BT2\text{-}ABS}_{\epsilon\to\epsilon}$.

Theorems (selected)

3. Meaning formula for $\mathsf{BPLUS}_{\epsilon\to\epsilon\to\epsilon}$
$\forall\, u_\epsilon\,.\,\forall\, v_\epsilon\,.\,((\mathsf{is\text{-}bnum}\; u_\epsilon \wedge \mathsf{is\text{-}bnum}\; v_\epsilon) \supset$
$(\mathsf{is\text{-}bnum}\; (u_\epsilon\; \mathsf{BPLUS}\; v_\epsilon) \wedge \quad ([\![u_\epsilon\; \mathsf{BPLUS}\; v_\epsilon]\!]_\iota = [\![u_\epsilon]\!]_\iota + [\![v_\epsilon]\!]_\iota))).$

5 Study 2: Test Case Formalized in Agda

As our goal is to, in part, compare the global approach and the local approach, the formalization in Agda [15, 18] eschews the use of its reflection capabilities[2]. Thus this formalization replaces the global type ϵ (of $\mathrm{CTT}_{\mathrm{uqe}}$) by a *set* of inductive types, one for each of the biform theories. This is still reflection, just hand-rolled. We also need to express formulas in FOL (as syntax), thus we need a type for that as well. To be more modular, that is broken up as a type for propositional logic with equality over any ground language, and a type for first-order logic which builds on that. We display some illustrative samples here; the full code is available in appendix B of [4].

An abstract theory is modeled as a *record*, with fields for the new type (pronounced Set in Agda), the two constants, and the two axioms. The host logic is dependently typed, and so the axioms refer to the constants just defined. One is not a field, but a defined constant.

```
record BT₁ : Set₁ where
  field
    nat : Set₀
    Z : nat
    S : nat → nat
    S≠Z : ∀ x → ¬ (S x ≡ Z)
    inj : ∀ x y → S x ≡ S y → x ≡ y

  One : nat
  One = S Z
```

For simplicity, we will take the built-in type $\mathbb{N}$, defined as an inductive type, as the *syntax* for natural numbers, which is also the syntax associated to the theory BT_1. Whereas in $\mathrm{CTT}_{\mathrm{uqe}}$ there is a global evaluation, here we also need to define evaluation explicitly (a subscript is used to indicate which theory it belongs to).

[2] As of early 2017, there is no official publication describing these features outside of the Agda documentation, but see [20, 21]

```
⟦_⟧₁ : ℕ → nat
⟦ 0 ⟧₁ = Z
⟦ suc x ⟧₁ = S ⟦ x ⟧₁
```

The accompanying code furthermore proves some basic coherence theorems which are elided here. We make two further definitions (GroundLanguage describing some language features, and FOL as our definition of first order logic) which will be explained in more detail on the next page.

```
nat-lang : GroundLanguage nat
nat-lang = record { Lang = λ X → ℕX (Carrier X)
                  ; value = λ {V} → val {V} }
  where
    val : {V : DT} → ℕX (Carrier V) → (Carrier V → nat) → nat
    val z env = Z
    val {V} (s e) env = S (val {V} e env)
    val (v x) env = env x

module fo₁ = FOL nat-lang
```

We can also demonstrate that the natural numbers are a model:

```
ℕ-is-T1 : BT₁
ℕ-is-T1 = record { nat = ℕ ; Z = 0 ; S = suc
  ; S≠Z = λ x → λ () ; inj = λ { x .x refl → refl } }
```

One of the languages needed is an extension of the naturals which allows variables:

```
data ℕX (Var : Set₀) : Set₀ where
  z : ℕX Var
  s : ℕX Var → ℕX Var
  v : Var → ℕX Var
```

But where the informal description in section 3 can get away with saying "Logic: FOL" and "Transformers: Recognizers for the formulas of the theory", here we need to be very explicit. To do so, we need to define some language infrastructure.

One of the important concepts is that of a *language with variables*, in other words a language with a reasonable definition of substitution. This requires *variables* to come from a type which has the structure of a decidable setoid (from the Agda library DecSetoid, and denoted DT below).

A language, expressed as an inductive type, is closed, i.e., cannot be extended. If a language does not have variables, we cannot add them. One solution is to deal with *contexts* as first-class citizens. While that is likely the best long-term solution, here we have gone with something simpler: create another language which does, and show that its variable-free fragment is equivalent to the original. As that aspect of our development is straightforward, albeit tedious, we elide it.

As we are concerned with statements in first-order logic over a variety of languages, it makes sense to modularize this aspect somewhat. Note that, as we are building syntax via inductive types, we can either build these as functors and

then use a fixpoint combinator to tie the knot, or we can just bite the bullet and make one large definition. For now, we chose the latter. We do parametrize over a *ground language with variables*. In turn, this is defined as a type parametrized by a decidable setoid along with an evaluation function into some type T.

```
record GroundLanguage (T : Set₀) : Set₁ where
  open DecSetoid using (Carrier)
  field
    Lang : DT → Set₀
    value : {V : DT} → Lang V → (Carrier V → T) → T
```

A logic over a language (with variables), is then also a parametrized type as well as a parametrized interpretation into types. The definition is almost the same, except that a ground language interprets into T and a logic into Set_0.

```
record LogicOverL (T : Set₀) (L : GroundLanguage T) : Set₁ where
  open DecSetoid using (Carrier)
  field
    Logic : DT → Set₀
    ⟦_⟧_ : ∀ {V} → Logic V → (Carrier V → T) → Set₀
```

The definition of first order logic is then straightforward.

```
module FOL {T : Set₀} (L : GroundLanguage T) where
  open DecSetoid using (Carrier)
  open GroundLanguage L

  data FOL (V : DT) : Set where
    tt : FOL V
    ff : FOL V
    _and_ : FOL V → FOL V → FOL V
    _or_ : FOL V → FOL V → FOL V
    not : FOL V → FOL V
    _⊃_ : FOL V → FOL V → FOL V
    _==_ : Lang V → Lang V → FOL V
    all : Carrier V → FOL V → FOL V
    exist : Carrier V → FOL V → FOL V

  override : {V : DT} → (Carrier V → T) → Carrier V → T → (Carrier V → T)
  override {V} f x t y with DecSetoid._≟_ V y x
  ... | yes _ = t
  ... | no _  = f y
```

We can also prove that FOL is a logic over L by providing an interpretation. Of course, as we are modeling classical logic into a constructive logic, we have to use a double-negation embedding. We also choose to interpret the logic's equality _==_ as *propositional equality*, but we could make that choice a parameter as well.

```
  LoL-FOL : LogicOverL T L
  LoL-FOL = record { Logic = FOL ; ⟦_⟧_ = interp }
    where
```

```
interp : {Var : DT} → FOL Var → (Carrier Var → T) → Set₀
interp tt env = ⊤
interp ff env = ⊥
interp (e and f) env = interp e env × interp f env
interp (e or f) env = ¬ ¬ (interp e env ⊎ interp f env)
interp (not e) env = ¬ (interp e env)
interp (e ⊃ f) env = (interp e env) → (interp f env)
interp (x == y) env = value x env ≡ value y env
interp {V} (all x p) env  = ∀ z → interp p (override {V} env x z)
interp {V} (exist x p) env = ¬ ¬ (Σ T (λ t → interp p (override {V} env x t)))
```

With the appropriate infrastructure in place, it is now possible to define BT_6 from the theories it extends.

```
record BT₆ {t₁ : BT₁} (t₂ : BT₂ t₁) (t₅ : BT₅ t₁) : Set₁ where
  open VarLangs using (XV; x)
  open DecSetoid using (Carrier)
  open BT₂ t₂ public
  open fo₂ using (FOL; tt; ff; LoL-FOL; _and_; all)
  open LogicOverL LoL-FOL

  field
    induct : (e : FOL XV) →
      ⟦ e ⟧ (λ { x → ⟦ 0 ⟧₁ }) →
      (∀ y → ⟦ e ⟧ (λ {x → y}) → ⟦ e ⟧ (λ {x → S y})) →
      ∀ y → ⟦ e ⟧ (λ {x → y})
  postulate
    decide : ∀ {W} → (Carrier W → nat) → FOL W → FOL NoVars
    meaning-decide : {W : DT} (env : Carrier W → nat) → (env' : ⊥ → nat) →
      (e : FOL W) →
      let res = decide env e in
      (res ≡ tt ⊎ res ≡ ff) × (⟦ e ⟧ env) ≃ (⟦ res ⟧ env')
```

While section 4 presents the *flattened* theory, here we need only define what is new over the extended theory, namely an induction schema, a decision procedure and its meaning formula.

Here is a guide to understanding the above definition: (1) XV is a (decidable) type with a single inhabitant, x. (2) All fields of BT_2 are made publicly visible for BT_6. (3) The language of first-order logic FOL over t_2 (and some of its constructors) is also made visible. (4) (λ {x → y}) denotes a substitution for the single variable x. (5) ≃ denotes *type equivalence*.

We represent numerals as vectors (of length at least 1) of binary digits.

```
data BinDigit : Set where zero one : BinDigit
data BNum : Set where
  bn : {n : ℕ} → Vec BinDigit (suc n) → BNum
```

This then allows a straightforward implementation of bplus to add numerals. It is then possible to *prove* that the meaning function for bplus is a theorem.

```
bplus-is-+ : ∀ x y → ⟦ bplus x y ⟧₂ ≡ ⟦ x ⟧₂ + ⟦ y ⟧₂
```

6 Comparison of the Two Formalizations

As expected, we were able to formalize this network of theories using both methods. Neither are fully complete; both are missing the actual decision procedures (which would be large undertakings). In particular,

- The CTT_{uqe} formalization is missing the definition of the language recognizers, as well as the full assurance of being mechanically checked. It has no "implementation" of any transformers.
- The Agda version implements evaluation but not substitution — which means that the induction statement in BT5–BT7 are not quite the same as in CTT_{uqe}; the models will be the same however. It also does not implement any theory morphisms, as record definitions are not first-class in Agda.

More importantly, because of our (explicit) choice to contrast the global and local approaches, each version uses different infrastructure to reason about syntax.

- CTT_{uqe} has a built-in inductive type of "all syntax", along with quotation and evaluation operators for the entire language of expressions.
- In the local approach, a new inductive type for each new "language" (the numerals, the numerals with plus, the numerals with plus and times, all three of these augmented with variables, first order logic, binary digits, binary numerals) has to be created. For many of these, a variety of traversals (folds) have to be implemented "by hand" even though the recursion patterns are obvious, at least to humans. Some of these are evaluation operators (one per language). There is no formal quotation operator.

The Agda version has a number of extra features: some transformers (such as for bplus and btimes) are implemented. Furthermore, the *meaning formula* for bplus is shown to be a theorem. A variety of coherence theorems are also shown, to gain confidence that the theories really are the ones we want.

It is worth remarking that defining the language of first-order formulas is complicated in *both* versions. This has been noticed before by people doing programming language meta-theory with proof assistants: encoding languages, especially languages with binders (such as FOL) along with traversals and basic reasoning can be very tedious [1].

The most notable differences in the two formalizations are:

1. Because FOL is classical, but Agda's host logic is constructive, a double-negation embedding was needed.
2. The use of *type equivalence* instead of boolean equality for verifying that the interpretation of a formula of FOL and the results of the decision procedure are "the same",

3. Borrowing the notion of *contractibility* from HoTT [19], to encode *definite description*.
4. Extending the decision procedure to *closeable* terms (by providing an explicit, total valuation) instead of restricting to closed terms.

The first is basically forced upon us by Agda: it has no Prop type (unlike Coq), and so we do not know a priori that all interpretations of first-order formulas are actually 0-types. The second is an active design decision: the infrastructure required to define the meaning of *closed* which is useful in a constructive setting is quite complex[3]. We believe the third is novel. The fourth point requires deeper investigating.

7 Conclusion

We have proposed a biform theory graph test case composed of eight theories that encode natural number arithmetic and include a variety of useful transformers. We have formalized this test case (as a set of biform theories and theory morphisms) in CTT_{uqe} using the global approach (for metareasoning with reflection) and in Agda using the local approach. In both cases, we have produced substantial partial formalizations that indicate that full formalizations could be obtained with additional work.

Our results show that, by providing a built-in global infrastructure, the global approach has a significant advantage over the local approach. The local approach is burdened by the necessity to define an infrastructure — consisting of an inductive type and an evaluation operator for the type — for every set of expressions manipulated by a transformer. In general, new local infrastructures must be created each time a new theory is added to the theory graph. On the other hand, the global approach employs an infrastructure — consisting of an inductive type, a quotation operator, and an evaluation operator — for the entire set of expressions in the logic. This single infrastructure is used for every theory in the theory graph.

We recommend that future research is directed toward making the global approach for metareasoning with reflection into a practical method for formalizing biform theories. This can be done by developing and implementing logics such as CTT_{qe} [7, 8] and CTT_{uqe} [9] and by adding global infrastructures to proof systems such as Agda and Coq (see [20, 21] for work in this direction).

References

1. B. E. Aydemir, A. Bohannon, M. Fairbairn, J. N. Foster, B. C. Pierce, P. Sewell, D. Vytiniotis, G. Washburn, S. Weirich, and S. Zdancewic. Mechanized metatheory for the masses: The PoplMark challenge. In *Theorem Proving in Higher Order Logics*, volume 3603 of *Lecture Notes in Computer Science*, pages 50–65. Springer Berlin Heidelberg, 2005.

[3] It would require us to define *paths* in terms, bound and free variables along paths, quantification over paths, etc.

2. J. Barwise and J. Seligman. *Information Flow: The Logic of Distributed Systems*, volume 44 of *Tracts in Computer Science*. Cambridge University Press, 1997.
3. J. Carette and W. M. Farmer. High-level theories. In A. Autexier, J. Campbell, J. Rubio, M. Suzuki, and F. Wiedijk, editors, *Intelligent Computer Mathematics*, volume 5144 of *Lecture Notes in Computer Science*, pages 232–245. Springer, 2008.
4. J. Carette and W. M. Farmer. Formalizing mathematical knowledge as a biform theory graph: A case study. *Computing Research Repository (CoRR)*, abs/1704.02253 (42 pp.), 2017.
5. W. M. Farmer. Biform theories in Chiron. In M. Kauers, M. Kerber, R. R. Miner, and W. Windsteiger, editors, *Towards Mechanized Mathematical Assistants*, volume 4573 of *Lecture Notes in Computer Science*, pages 66–79. Springer, 2007.
6. W. M. Farmer. The formalization of syntax-based mathematical algorithms using quotation and evaluation. In J. Carette, D. Aspinall, C. Lange, P. Sojka, and W. Windsteiger, editors, *Intelligent Computer Mathematics*, volume 7961 of *Lecture Notes in Computer Science*, pages 35–50. Springer, 2013.
7. W. M. Farmer. Incorporating quotation and evaluation into church's type theory. *Computing Research Repository (CoRR)*, abs/1612.02785 (72 pp.), 2016.
8. W. M. Farmer. Incorporating quotation and evaluation into Church's type theory: Syntax and semantics. In M. Kohlhase, M. Johansson, B. Miller, L. de Moura, and F. Tompa, editors, *Intelligent Computer Mathematics*, volume 9791 of *Lecture Notes in Computer Science*, pages 83–98. Springer, 2016.
9. W. M. Farmer. Theory morphisms in Church's Type theory with quotation and evaluation. *Computing Research Repository (CoRR)*, abs/1703.02117 (15 pp.), 2017.
10. W. M. Farmer, J. D. Guttman, and F. J. Thayer. Little theories. In D. Kapur, editor, *Automated Deduction—CADE-11*, volume 607 of *Lecture Notes in Computer Science*, pages 567–581. Springer, 1992.
11. W. M. Farmer, J. D. Guttman, and F. J. Thayer. IMPS: An Interactive Mathematical Proof System. *Journal of Automated Reasoning*, 11:213–248, 1993.
12. R. D. Jenks and R. S. Sutor. *Axiom : The Scientific Computation System*. Springer, 1992.
13. M. Kohlhase. Mathematical knowledge management: Transcending the one-brain-barrier with theory graphs. *European Mathematical Society (EMS) Newsletter*, pages 22—-27, June 2014.
14. U. Norell. *Towards a Practical Programming Language based on Dependent Type Theory*. PhD thesis, Chalmers University of Technology, 2007.
15. U. Norell. Dependently typed programming in Agda. In A. Kennedy and A. Ahmed, editors, *Proceedings of TLDI'09*, pages 1–2. ACM, 2009.
16. F. Rabe and M. Kohlhase. A scalable model system. *Information and Computation*, 230:1–54, 2013.
17. C. Smoryński. *Logical Number Theory I: An Introduction*. Springer, 1991.
18. Agda Team. Agda wiki. `http://wiki.portal.chalmers.se/agda/pmwiki.php`. Accessed: 2017-05-15.
19. The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. `https://homotopytypetheory.org/book`, Institute for Advanced Study, 2013.
20. P. van der Walt. Reflection in Agda. Master's thesis, Universiteit Utrecht, 2012. `https://dspace.library.uu.nl/handle/1874/256628`.
21. P. van der Walt and W. Swierstra. Engineering proof by reflection in Agda. In R. Hinze, editor, *Implementation and Application of Functional Languages*, volume 8241 of *Lecture Notes in Computer Science*, pages 157–173. Springer, 2012.