Cardinal Arithmetic in Weak Theories

Albert Visser

Department of Philosophy, Utrecht University

Heidelberglaan 8, 3584 CS Utrecht, The Netherlands

email: `Albert.Visser@phil.uu.nl`

March 22, 2008

Abstract

In this paper we develop the theory of cardinals in the theory COPY. This is the theory of two total, jointly injective binary predicates in a second order version, where we may quantify over binary relations. The only second order axioms of the theory are the axiom asserting the existence of an empty relation and the adjunction axiom, which says that we may enrich any relation R with a pair x, y. The theory COPY is strictly weaker than the theory AS, adjunctive set theory. The relevant notion of *weaker* here is *direct interpretability*. We will explain and motivate this notion in the paper. A consequence is that our development of cardinals is inherited by stronger theories like AS. We will show that the cardinals satisfy (at least) Robinson's Arithmetic Q. A curious aspect of our approach is that we develop cardinal multiplication using neither recursion nor pairing, thus diverging both from Frege's paradigm and from the tradition in set theory. Our development directly uses the universal property characterizing the product that is familiar from category theory.

The broader context of this paper is the study of a double degree structure: the degrees of (relative) interpretability and the finer degrees of direct interpretability. Most of the theories studied are in one of two degrees of interpretability: the bottom degree of predicate logic or the degree of Q. The theories will differ significantly if we compare them using direct interpretability.

Key words: interpretations, weak arithmetical theories, adjunctive set theory, cardinals

MSC2000 codes: 03B30, 03F25, 03F30

1 Introduction

There is a wide and natural class of theories called 'weak theories'. These are the theories that are mutually interpretable with Robinson's Arithmetic Q.[1]

[1] The notion of weak theory is informal in the literature. It is not fully clear which theories are 'weak'. Thus, another plausible explication would be: a theory is weak if it is *locally* mutually interpretable with Q.

Examples of weak theories are the following.

a. Arithmetical theories like Robinson's Arithmetic Q, $I\Delta_0$, Buss's theory S^1_2 and $I\Delta_0 + \Omega_1$.

b. Various theories of concatenation like Szmielew & Tarski's theory F (see: [TMR53], p86) and Grzegorczyk's theory TC (see: [Grz05]).

c. Theories of sequences, like the one introduced by Pavel Pudlák in [Pud83].

d. Various weak set theories like adjunctive set theory AS, a set theory introduced by Pudlák in [Pud83] (see also: [MPS90]).[2] Pudlák's theory is, modulo some inessential details, the same as a weak set theory introduced and studied by Tarski & Szmielew, *minus extensionality*. See [ST50] and [TMR53].

All these theories are mutually interpretable with each other. This insight can be viewed as a basic theorem of what Harvey Friedman calls *strict reverse mathematics*, where one classifies mathematically natural weak systems.

There is a strong feeling that the theories mutually interpretable with Q form a natural class. For one thing, in the light of results of Solovay and Pudlák, the theories in this degree seem to be the minimal natural theories in which we can do formalization of syntax in the care free way, so that we can give the usual proof of the Second Incompleteness Theorem —at least when we choose the natural numbers of the theory wisely. This feeling of naturality suggests that Q should be characterizable by some perspicuous (universal) property. Regrettably, no such property has been found (yet). For example, one would hope that the degrees of interpretability of finitely axiomatized, essentially undecidable theories have a unique minimal element given by Q. This is not so.[3]

Modulo mutual interpretability all the theories of the groups (a), (b), (c), (d) are the same, but, still, there is an important division between (a) and (b) on the one hand and (c) and (d) on the other. To explain why this is so, let us make a slight detour. Consider an arbitrary theory U in predicate logic. What would it mean to say that this theory 'has pairing' or 'supports pairing'? Well, it means that there is a formula $Axyz$ in the language of U such that U proves that, for every x and y, there is a z such that $Axyz$, and that, for every x, y, x', y', z, if $Axyz$ and $Ax'y'z$, then $x = x'$ and $y = y'$. We can reformulate this definition in terms of interpretability as follows. We can define a theory of pairing with a basic ternary relation pair that has as axioms that for every x and y, there is a z such that $\mathsf{pair}(x, y, z)$, and that, for every x, y, x', y', z, if $\mathsf{pair}(x, y, z)$ and $\mathsf{pair}(x', y', z)$, then $x = x'$ and $y = y'$. We now would like to say: a theory has pairing if it interprets the theory of pairing. This is however not quite right. In fact every theory interprets our theory of pairing in the sense of relative interpretability where identity need not necessarily be translated to identity. Our theory of pairing has a one point model and any theory with a one

[2] A nice variant of AS was given by Harvey Friedman in his Tarski Lecture *Interpretations, according to Tarski*.

[3] This follows from a result of William Hanf. See [Han65].

point model is interpretable in any other theory. What went wrong is of course that we should neither allow a change of our domain nor a change of the identity relation. We call an interpretation *direct* if it is unrelativized and if it sends identity to identity. Now we can say: U has pairs iff it directly interprets the theory of pairing. We can repeat this idea for certain basic theories of sequences and sets. Thus, we can say that an abritrary theory U has sets iff it directly interprets the basic theory, AS, of sets. The theories classified under (b) and (c) are examples of theories that can play the role of basic theories of a certain *kind of container*. It turns out that the basic theories under (b) and (c) are mutually directly interpretable. This implies that if an arbitrary theory 'has sets' then it also 'has sequences' and vice versa. Theories that interpret adjunctive se theory AS will be called *sequential*.

Sequential theories were independently introduced by Pavel Pudlák and Harvey Friedman, who called them *adequate theories*. See, e.g., [Pud83], [Pud85], [Smo85], [MPS90]. For a survey, see [Vis07].

An important ingredient of the usefulness of a theory like AS is the fact that it (relatively) interprets Robinson's Arithmetic Q. In this paper we zoom in on the question of the interpretability of Q in weak container theories. Here is a brief history of the result that Q is interpretable in the salient weak container theory Adjunctive Set Theory or AS.

1. In the paper [ST50], Wanda Szmielew and Alfred Tarski announce the interpretability of Q in a theory S that is essentially AS *plus extensionality*.[4] See also [TMR53], p34. No proof was published.

2. A proof of the Szmielew-Tarski result is given by George Collins and Joseph Halpern in [CH70]. Collins and Harper did not have Solovay's method of shortening cuts available.[5] So, it is rather amazing that they manage to prove addition and multiplication total. They succeed by a clever choice of values for plus and times whenever the recursive definition does not turn out a value. Their interpretation of Q is direct.

3. Franco Montagna and Antonella Mancini, in their paper [MM94], give an improvement of the Szmielew-Tarski result. They prove that Q can be interpreted in an extension N of AS in which we have the functionality of empty set and the operation of adjoining of singletons. They sketch a proof of the Herbrand consistency of their set theory that can be proved in a predicative arithmetic.

4. In appendix III of [MPS90], Jan Mycielski, Pavel Pudlák and Alan Stern provide the ingredients of the interpretation of Q in AS.[6] They do not develop the theory of addition and multiplication, but these can be treated in familiar ways using the theory of sequences that provided by their argument. See e.g. [Pud83] or [HP91].

[4] John Burgess in [Bur05] , p90-91, calls this theory ST, for Szmielew-Tarski set theory.

[5] Solovay's method dates from roughly 1976. See the unpublished letter [Solle].

[6] Mycielski, Pudlák and Stern do not provide a name for their weak set theory. They call any theory that directly interprets AS: a weak set theory.

5. John Burgess in his [Bur05], Section 2.2, provides a variant of the Montagna-Mancini argument.

The style of the interpretation of Q provided in each of these proofs is by treating the numbers mainly as finite ordinals. Specifically, the main part of the definitions of addition and multiplication consists of a recursion. We will provide yet another proof of the interpretability of Q in AS.

Why is the interpretability of Q in AS interesting? A first reason is that it plays an important role in verifying that adjunctive set theory directly interprets a theory of sequences. The role of this insight is as follows. The use of sequential theories has two sides. The first is verification of the fact that a given theory is sequential. This should be *as easy as possible*. For this purpose, we only have to verify that AS is directly interpretable in the given theory. the theory AS is ideal with its simple signature and modest axioms. The second side is application of the fact that a theory is sequential. For this it is essential that we have a *rich* theory of sequences and coding available. To move from the verification side to the application side, we have a bridging theorem that says that a rich theory of numbers and sequences is directly interpretable in the poor theory AS. The key part of this bridging theorem is to provide suitable numbers in AS. To have an interpretation that validates Q is sufficient here.

A second source of interest, is the desire to understand what is minimally involved in the connection between sets and numbers. The development of Q in AS (and, in the even weaker theory COPY) shows that we can proceed from fairly minimal assumptions.

What is to be found in the present paper? We provide an alternative proof of the theorem of Mycielski, Pudlák and Stern. The main difference in approach is that we develop Q as a theory of cardinals in the theory COPY that is directly interpretable in AS.[7]

In order to implement cardinals, the central problem is to develop the Cartesian product. One obvious difficulty is that we do not have a functional pairing in AS, so that the usual set theoretic definition of product does not yield a unique class of pairs. One can develop products in two ways.

The first is to mimic the usual development in set theory as much as possible, working with many isomorphic classes of pairs as 'cartesian products of the form $X \times Y$'. To realize such a development the appropriate theory to work with is a theory CART of pairs and classes, that is mutually directly interpretable with AS. The theory CART has non-functional pairing. Basically, following the first strategy, one would develop a certain *implementation* of the product. Such a development is undoubtedly possible, but one has to work hard to cope with the non-functionality of pairing and, e.g., to provide the necessary isomorphisms between any product of the form $(X \times Y) \times Z$ and any product of the form $X \times (Y \times Z)$.[8]

[7] The remark that Q can be viewed as a theory of cardinals is attributed by John Burgess to Saul Kripke. See [Bur05], p56.

[8] Associativity of multiplication is not an axiom of Q, but one needs it nevertheless to make multiplication a total function via Solovay style trickery.

The other way to define products is by their universal property and to create sufficiently many objects having this universal property. Here we do not work with a specific implementation. As bonus, we get the associativity and commutativity of multiplication for free. Moreover, we are able to interpret Q in a theory that is strictly weaker than AS modulo direct interpretability.[9] The theory COPY seems precisely the theory to work with since it allows us to construct the isomorphic copies of classes, needed to develop disjoint union. COPY is the theory of two jointly injective binary predicates enriched with a theory of binary relations containing just the existence-of-empty-relation axiom and the adjunction axiom for relations.

We opt for the second strategy. To attain our goal, we develop a category in COPY. I feel that this development holds some independent interest. We get a bit more than just the cardinals with addition and multiplication, e.g., in our category, we have all finite limits. Moreover, it seems to offer us some further possibilities for further research, like the development of exponentiation as a partial function in the set theoretic way.

The paper contains a few extras. We give a slow exposition and scenic tour of some theories 'in the environment of' AS. E.g., we show that AS is mutually directly interpretable with the theory of pairs and classes called CARD.

Remark 1.1 Burgess provides an interpretation of Q in the Montagna-Mancini variant of AS. He does this by showing that the extension of the successor axioms of Q with a second order part concerning binary relations, axiomatized by the existence-of-empty-relation axiom and the adjunction axiom for relations. Let us call this theory SUCC^+. We note the analogy of the role of COPY in our proof and of SUCC^+ in Burgess proof. Of course, Burgess' is development is Frege style: he uses recursion. Our development is recursion free. ❑

Finally, the paper illustrates a methodological point. In reverse mathematics notions of reducibility play a central role. There are various such notions like conservativity[10] and interpretability. In strict reverse mathematics, where we study really weak theories, interpretability is an obvious choice of reduction. In this paper we illustrate the importance of the notion of *direct interpretability* as a notion of reduction, not to replace relative interpretability, but to use in conjunction with it.

Prerequisites

All necessary techniques are developed in the paper. For some background to the project, see e.g. [HP91]. Of category theory we only use the elementary facts about sum and product. See [Mac71].

[9] We will show that COPY does not have pairing of its objects.

[10] Note that conservativity *presupposes* interpretability: formula classes like Π^0_2 are usually only present via a designated interpretation.

2 Theories and Interpretations

We consider theories in many-sorted first order predicate logic. We assume that these theories have *officially* relational signature. Unofficially, we use function symbols, but these can be eliminated using a well-known unwinding procedure. We will consider a notion of reduction called *direct interpretability*. This reduction relation holds between *pointed theories*, i.e., theories with a designated sort, $\mathfrak{o}$, 'the sort of objects'. The pointed sort always has identity.

We will consider *piecewise, more-dimensional, many-sorted, relative interpretations with parameters, where identity is not necessarily translated as identity.* Since, the presence of parameters, being piecewise and more-dimensionality only play a minor role, we will give a careful definition of *one-dimensional, many-sorted, relative interpretations without parameters, where identity is not necessarily translated as identity.* We will briefly indicate how to extend the framework to being piecewise, more-dimensionality and parameters. We recommend the reader to skip these subsections and return to them if needed.

2.1 Translations

To define an interpretation, we first need the notion of translation. Let Σ and Ξ be finite signatures for many-sorted predicate logic with finitely many sorts. We assume that the sorts are specified with the signature. We also assume that the designated sort is also given by the signature. A *relative translation* $\tau : \Sigma \to \Xi$ is given by a triple $\langle \sigma, \delta, F \rangle$. Here σ is a mapping of the Σ-sorts to the Ξ-sorts. The mapping δ assigns to every Σ-sort $\mathfrak{a}$ a Ξ-formula $\delta^{\mathfrak{a}}$ representing the *domain* for sort $\mathfrak{a}$ of the translation. We demand that $\delta^{\mathfrak{a}}$ contains at most a designated variable $v_0^{\sigma\mathfrak{a}}$ of sort $\sigma\mathfrak{a}$ free. The mapping F associates to each relation symbol R of Σ a Ξ-formula $F(R)$. The relation symbol R comes equipped a sequence $\vec{\mathfrak{a}}$ of sorts. We demand that $F(R)$ has at most the variables $v_i^{\sigma\mathfrak{a}_i}$ free. We translate Σ-formulas to Ξ-formulas as follows:

- $(R(y_0^{\mathfrak{a}_0}, \cdots, y_{n-1}^{\mathfrak{a}_{n-1}}))^\tau := F(R)(y_0^{\sigma\mathfrak{a}_0}, \cdots, y_{n-1}^{\sigma\mathfrak{a}_{n-1}})$.

  (We assume that some mechanism for α-conversion is built into our definition of substitution to avoid variable-clashes.)
- $(\cdot)^\tau$ commutes with the propositional connectives;
- $(\forall y^{\mathfrak{a}}\, A)^\tau := \forall y^{\sigma\mathfrak{a}}\, (\delta^{\mathfrak{a}}(y) \to A^\tau)$;
- $(\exists y^{\mathfrak{a}}\, A)^\tau := \exists y^{\sigma\mathfrak{a}}\, (\delta^{\mathfrak{a}}(y) \wedge A^\tau)$.

Suppose τ is $\langle \sigma, \delta, F \rangle$. Here are some convenient conventions and notations.

- We write δ_τ for δ and F_τ for F.
- We write R_τ for $F_\tau(R)$.
- We will always use '$=^{\mathfrak{a}}$' for the (optional) identity of a theory for sort $\mathfrak{a}$. In the context of translating, we will however switch to '$E^{\mathfrak{a}}$'.

- We write $\vec{x} : \delta^{\vec{\mathfrak{a}}}$ for: $\delta^{\mathfrak{a}_0}(x_0^{\sigma\mathfrak{a}_0}) \wedge \ldots \wedge \delta^{\mathfrak{a}_{n-1}}(x_{n-1}^{\sigma\mathfrak{a}_{n-1}})$.
- We write $\forall\vec{x} : \delta^{\vec{\mathfrak{a}}}\ A$ for: $\forall x_0^{\sigma\mathfrak{a}_0} \ldots \forall x_{n-1}^{\sigma\mathfrak{a}_{n-1}}\ (\vec{x}{:}\delta^{\vec{\mathfrak{a}}} \to A)$.

  Similarly for the existential case.
- Suppose U is a theory of signature Ξ. Suppose further that $U \vdash \exists v\ \delta^{\mathfrak{a}}_{\tau} v$, for every Σ-sort $\mathfrak{a}$. We define: $\tau^{-1}(U) := \{A{\in}\mathsf{sent}_{\Sigma} \mid U \vdash A^{\tau}\}$.

2.2 Interpretations and Interpretability

A translation τ supports a *relative interpretation* of a theory U in a theory V, if, for all axioms A of U, we have $U \vdash A \Rightarrow V \vdash A^{\tau}$. (Note that this automatically takes care of the theory of identity. Moreover, it follows that $V \vdash \exists v_0\ \delta^{\mathfrak{a}}_{\tau} v_0$, for each Σ-sort $\mathfrak{a}$.) Thus, an interpretation has the form: $K = \langle U, \tau, V\rangle$.

Par abus de langage, we write 'δ_K' for: δ_{τ_K}; 'P_K' for: P_{τ_K}; 'A^K' for: A^{τ_K}, etc. We define:

- We write $K : U \lhd V$ or $K : V \rhd U$, for: K is an interpretation of the form $\langle U, \tau, V\rangle$.
- $V \rhd U :\Leftrightarrow U \lhd V :\Leftrightarrow \exists K\ K : U \lhd V$.

  We read $U \lhd V$ as: U is interpretable in V. We read $V \rhd U$ as: V interprets U.

2.3 Direct Interpretations

We now consider pointed theories, i.e., theories with a designated sort of objects $\mathfrak{o}$. An translation τ is *direct* if σ_{τ} preserves the designated sort, and, for the designated sort, τ is unrelativized and has absolute identity, i.e.:

- $\delta^{\mathfrak{o}}(v_0) :\leftrightarrow (v_0 = v_0)$,
- $v_0 E^{\mathfrak{o}}_{\tau} v_1 :\leftrightarrow v_0 = v_1$.

An interpretation is direct if it is based on a direct translation. We write $V \rhd_{\mathsf{dir}} U$, for V directly interprets U, etc. Direct interpretation is our main tool in this paper: we want to 'refine' our sort of classes without tinkering with our objects.

We will use the sum $T \boxplus U$ of pointed theories. This sum is obtained as follows. First we make the sorts and predicates of the theories disjoint *except the designated sort* $\mathfrak{o}$ *and except for the identity predicates for the designated sorts*. Then, we take the union of the modified theories.

Our sum is the sum in the category of direct interpretations. Thus, the sum is a bifunctor w.r.t. the preorder of direct interpretability, i.e., if $V \rhd_{\mathsf{dir}} U$ and $V' \rhd_{\mathsf{dir}} U'$, then $(V \boxplus V') \rhd_{\mathsf{dir}} (U \boxplus U')$.

The one-sorted theory of pure identity $\mathsf{identity}$ is in the lowest degree of direct interpretability together with any many-sorted predicate logic. We clearly have $\mathsf{identity} \boxplus U = U$.

We will speak about the degrees of direct interpretability inside a given degree of interpretability. Note that strictly speaking the switch from direct interpretability to ordinary interpretability involves removing the designation of the designated sort from the theory. We will ignore this subtlety in the text.

2.4 Expansion and Definitional Extension

A translation $\tau : \Sigma \to \Xi$ is an *expansion* if:

1. Σ has the same sorts as Ξ;
2. for every sort $\mathfrak{b}$ of Σ, $\sigma_\tau \mathfrak{b} = \mathfrak{b}$ and $\delta_\tau^{\mathfrak{b}} v_0 :\leftrightarrow v_0 = v_0$;
3. every predicate P of Ξ is also in Σ, and $P_\tau \vec{v} :\leftrightarrow P\vec{v}$.

Suppose τ is an expansion and U is a theory of signature Ξ. We call $\tau^{-1}U$ a *definitional extension* of U. Note that $\tau^{-1}U$ is axiomatizable over U by axioms of the form $\vdash \forall \vec{v}\,(P\vec{v} \leftrightarrow P_\tau \vec{v})$.[11]

2.5 Local Interpretability

We say that a (pointed) theory V locally (directly) interprets a (pointed) theory U if, for any finite subtheory U_0 of U, we have $V \rhd_{(\mathsf{dir})} U_0$. We write $V \rhd_{(\mathsf{dir},)\mathsf{loc}} U$ for: V locally (directly) interprets U.

2.6 Multidimensionality

We can extend the notion of interpretation to the case of multidimensional interpretations by sending a sort $\mathfrak{a}$ of the interpreted theory via σ to *a sequence of sorts* of the interpreting theory. A domain formula $\delta^{\mathfrak{a}}$ will have a sequence of variables as arguments. These variables have the sorts given by $\sigma(\mathfrak{a})$. If we are considering pointed theories and direct interpretations, the designated sort $\mathfrak{o}$ goes to the singleton sequence $\langle \mathfrak{o} \rangle$. Etc.

2.7 Parameters

We can extend the notion of interpretation to the case of interpretation with parameters by allowing extra parameters from a given finite sequence $\vec{w}$ into the δ_τ and P_τ. As extra data we need a fixed formula $A\vec{w}$ of the language of the interpreting theory representing the intended range of the parameters. E.g., in the Poincaré disk interpretation of the hyperbolic plane in the Euclidean plane, the parameters could be w_0, w_1, where w_0 is the centre of the Poincaré disk and where w_0 is a point on the circumference of the disk. The formula A would be

[11] Definitional extensions are definitionally equivalent with their original theories. This means that they are isomorphic in an appropriate category of interpretations with their original theories. We will follow the usual practice of confusing a theory with its definitional extensions.
There should also be a good notion of definitional extension for the case where we add new sorts. We postpone developing this to another paper.

$w_0 \neq w_1$. We define: $\langle U, A\vec{w}, \tau\vec{w}, V\rangle$ is an interpretation iff $V \vdash \exists\vec{w}\ A\vec{w}$ and, for any axiom B of U, we have $V \vdash \forall\vec{w}\ (A\vec{w} \to B^{\tau\vec{w}})$.

2.8 Piecewise Interpretations

The idea of piecewise interpretations[12] is that we can develop the domains for each sort as built up from a finite number of pieces. These pieces may overlap. However the same original object may pose as two different objects depending on which piece we are considering.

For example, let us define a one dimensional, one sorted, parameter free translation with two pieces, for a theory with just one binary predicate P. Our translation τ provides two domain pieces δ_0 and δ_1. The δ_i are formulas of the target language containing just v_0 free. We have a function F that assigns formulas $A_{ij}(v_0, v_1)$ to P and 0,1-pairs ij. Similarly for the identity. Our translation function has as inputs formulas B of source language and assignments α of 0 or 1 to the free variables of B. We give the clauses for Pxy, conjunction and universal quantification.

- $(Pxy)^{\tau,\alpha} := A_{\alpha(x),\alpha(y)}(x, y)$, where α has domain $\{x, y\}$;
- $(B \wedge C)^{\tau,\alpha} := B^{\tau,\alpha\restriction\mathsf{FV}(B)} \wedge C^{\tau,\alpha\restriction\mathsf{FV}(C)}$, where α has domain $\mathsf{FV}(B \wedge C)$;
- $(\forall x\ B)^{\tau,\alpha} := \forall x{:}\delta_0\,(B)^{\tau,\alpha\cup\{\langle x,0\rangle\}} \wedge \forall x{:}\delta_1\,(B)^{\tau,\alpha\cup\{\langle x,1\rangle\}}$, where α has domain $\mathsf{FV}(\forall x\ B)$.

Note that τ will be the empty assignment for sentences. The rest of the development is as expected. Note that any theory with a finite model has a piecewise, one-dimensional, parameter free interpretation in $\mathsf{identity}$, the one-sorted theory of pure identity.

A natural example of a piecewise interpretation is 'adding the unit to the theory of semigroups'.

Remark 2.1 Prima facie, piecewise translations are rather costly. Translations that are not piecewise yield p-time transformations of formulas (if we handle the needed alpha-conversions in a sufficiently smart way). Piecewise translations, on the other hand may be exponential.

Fortunately, in rather general circumstances piecewise translations can be eliminated. E.g., if the target theory T, i.e., the interpreting theory T, does not have a one-element model, then a piecewise translation can always be replaced by a a multi-dimensional interpretation with parameters that is 'the same' in the sense that there is a T-definable, T-verifiable isomorphism between these translations (considered as inner model constructions). ❑

[12] I learned the idea of piecewise interpretability in a slightly less general form from Harvey Friedman.

3 Boolean Operations on Classes

In this section, we study the construction of Boolean operations in a very weak theory of classes.

The theory acl is a two-sorted theory, with a sort of objects $\mathfrak{o}$ and a sort of classes or concepts $\mathfrak{c}$. $x, y, z, \ldots$ will range over objects and $X, Y, Z, \ldots$ will range over classes. Our predicates are identity for the object sort and $\in$ of type $\mathfrak{oc}$. Our axioms are as follows.

acl1. $\vdash \exists X\, \forall x\; x \notin X$,

acl2. $\vdash \forall X, x\, \exists Y\, \forall y\; (y \in Y \leftrightarrow (y \in X \vee y = x))$.

We will 'load' acl modulo *mutual direct interpretability* to a (seemingly) much stronger theory.

We extend acl with an identity relation on classes, setting:

$$X =_{\mathfrak{c}} Y :\leftrightarrow \forall z\; (z \in X \leftrightarrow z \in Y).$$

This justifies the introduction of a constant $\emptyset$, a function $x \mapsto \{x\}$ and partial functions $\cup$ and $\cap$. Note that $\emptyset$ will be provably inhabited and that the mapping $X, x \mapsto X \cup \{x\}$, will be provably total.

Remark 3.1 The theory acl is very weak, because it has a model with one object and two classes. Thus it will be (piecewise) interpretable in $\mathsf{identity}$, the one-sorted theory of pure identity. Thus, $\mathsf{identity}$ interprets acl extended with full impredicative comprehension —since full comprehension is trivial on a domain with one object.

The remark illustrates the importance of restricting oneself to *direct* interpretations when studying acl. Corollary 4.4 tells us that acl is not in the minimal degree of direct interpretability. We will provide a theory arel that has a finite model, that does directly interpret acl, but that is not directly interpretable in acl. See Theorem 5.1. Thus, there are at least three strictly ascending degrees of direct interpretability in the minimal degree of interpretability. ❑

Let class_0 be the following predicate of classes:

- $\mathsf{class}_0(X) :\leftrightarrow \forall Y\; (Y \cup X) \downarrow$.

We have the following lemma.

Lemma 3.1 (acl) *The predicate* class_0 *is closed under empty class, singletons and union.*

Proof

It is easy to see that class_0 contains the empty class and all singletons. Let X and X' be in class_0. Then $X \cup X'$ exists. Moreover, for any Y, $(Y \cup X) \cup X'$ exists, and this is equal to $Y \cup (X \cup X')$. ❑

By restricting our domain of classes to class_0, we obtain a direct interpretation of $\mathsf{acl}_1 := \mathsf{acl} + \mathsf{union}$, where union is the axiom stating that union is total.

We define the predicate class_1 as follows.

- $\mathsf{class}_1(X) :\leftrightarrow \forall Y\ (Y \cap X)\downarrow$.

Lemma 3.2 (acl_1) *The predicate* class_1 *is closed under empty class, singletons, union and intersection.*

Proof

It is clear that class_1 contains the empty class. Note that $Y \cap \{x\}$ is $\emptyset$ if $x \notin Y$, and is $\{x\}$ if $x \in Y$. In both cases $Y \cap \{x\}$ exists. Suppose X and X' are in class_1. Let Y be arbitrary. Then, $Y \cap (X \cup X') = (Y \cap X) \cup (Y \cap X')$ and $Y \cap (X \cap X') = (Y \cap X) \cap X'$. So, $Y \cap (X \cup X')$ and $Y \cap (X \cap X')$ exist. We may conclude that $X \cup X'$ and $X \cap X'$ are in class_1. ❑

Restricting to class_1 gives us a direct interpretation of $\mathsf{acl}_2 := \mathsf{acl}_1 + \mathsf{intersection}$, where $\mathsf{intersection}$ is the axiom stating that intersection is total. We proceed to work in $\mathsf{acl}_2 \boxplus U$, where U is any theory. For any definable predicate $\mathcal{P}$ of classes we take:

- $\mathcal{P}^{\mathsf{s}} X :\leftrightarrow \forall Y{\subseteq}X\ \mathcal{P}Y$.

Lemma 3.3 ($\mathsf{acl}_2 \boxplus U$) *Suppose $\mathcal{P}$ is a predicate of classes which is closed under empty class, singletons, and union. Then, $\mathcal{P}^{\mathsf{s}}$ is closed under empty set, singletons and union. Moreover, $\mathcal{P}^{\mathsf{s}}$ is downwards closed w.r.t. $\subseteq$, and, hence, closed under intersection.*

Proof

Clearly, the empty class and all singletons are in $\mathcal{P}^{\mathsf{s}}$. Downwards closure w.r.t. $\subseteq$ is trivial. Suppose X and X' are in $\mathcal{P}^{\mathsf{s}}$ and $Y \subseteq X \cup X'$. We have: $Y \cap X \subseteq X$ and $Y \cap X' \subseteq X'$. So, $Y \cap X$ and $Y \cap X'$ are in $\mathcal{P}^{\mathsf{s}}$ and, hence, in $\mathcal{P}$. It follows that $Y = (Y \cap X) \cup (Y \cap X')$ is in $\mathcal{P}$. ❑

Consider, in $\mathsf{acl}_2 \boxplus U$, any predicate $Ax\vec{y}\vec{Y}$. Now consider the predicate class_A of all Z such that, for all $\vec{y}, \vec{Y}$, we have that $\{z{\in}Z \mid Ax\vec{y}\vec{Y}\}$ exists. It is easily seen that class_A is closed under empty class, singletons and union, and that it is downwards closed w.r.t. $\subseteq$: if $\{z{\in}Z \mid Az\vec{y}\vec{Y}\}$ exists and $W \subseteq Z$, then:

$$\{w{\in}W \mid Aw\vec{y}\vec{Y}\} = \{z{\in}Z \mid Az\vec{y}\vec{Y}\} \cap W.$$

Thus, we can get Aussonderung for any A, when we restrict ourselves to forming subclasses of classes from class_A:

- $\mathsf{acl} \boxplus U \vdash \forall Z : \mathsf{class}_A\ \{z{\in}Z \mid Az\}$ exists .

Note that $\{z{\in}Z \mid Axy\vec{y}\vec{Y}\}$ will be itself in class_A. This principle, however, is not absolute. If we relativize to a predicate that is closed under empty class, singletons, union and downwards closed w.r.t. $\subseteq$, the principle is not preserved —since the meaning of A changes. We can remedy this by restricting ourselves to $\Delta_0^{\subseteq}(U)$-formulas A, i.e. $(\mathsf{acl} \boxplus U)$-formulas where all class-quantifiers are $\subseteq$-bounded. Thus, we get:

Theorem 3.2 *The theory* $\mathsf{acl}_3(U) := (\mathsf{acl}_1 \boxplus U) + \Delta_0^{\subseteq}(U)$-Aussonderung *is locally directly interpretable in* $\mathsf{acl} \boxplus U$*. This theory is preserved to relativizations of the classes to subdomains that are closed under empty class, singletons, union and that are downwards closed w.r.t.* $\subseteq$*.*

Remark 3.3 Our second order variables have two 'standard meanings': the strong one, where the variables range over arbitrary sets of objects, and the weak one where the variables range over finite sets of objects. thus, we have strong standard models and weak standard models. Note that our interpretations all behave like the identity interpretation both in weak and in strong standard models. ❑

4 Adjunctive Set Theory

The language of the one-sorted theory AS has, apart from identity, just one binary predicate $\in$. The theory is given by the following axioms.

AS1. $\vdash \exists x\, \forall y\; x \notin y$,

AS2. $\vdash \forall x, y\, \exists z\, \forall u\; (u \in z \leftrightarrow (u \in x \vee u = y))$.

There is a slight variant AS^{s}, where we have an additional unary predicate set, which is axiomatized by the following axioms.

AS^{s}1. $\vdash \exists x{:}\mathsf{set}\, \forall y\; x \notin y$,

AS^{s}2. $\vdash \forall x{:}\mathsf{set}\, \forall y\, \exists z{:}\mathsf{set}\, \forall u\; (u \in z \leftrightarrow (u \in x \vee u = y))$,

AS^{s}3. $\vdash \forall x, y\; (x \in y \to y : \mathsf{set})$.

One easily shows that $\mathsf{AS} \equiv_{\mathsf{dir}} \mathsf{AS}^{\mathsf{s}}$. The theory AS has the advantage of economy of signature. On the other hand, for our general program, it seems more natural to distinguish urelements from empty sets as in AS^{s}. Since we are thinking mainly modulo mutual direct interpretability, we will use both versions as convenience dictates.

We give two possible ways of linking AS to acl. The first is via the Frege relation. The second is by adding pairing to acl.

4.1 Adjunctive Set Theory & the Frege Relation

We add to the language of acl the binary Frege relation $\succ$ of type $\mathfrak{co}$. The theory $\mathsf{acl}^{\mathsf{fr}}$ is axiomatized by.

$\mathsf{acl}^{\mathsf{fr}}1.$ $\vdash \exists X\, \forall x\; x \notin X$,

$\mathsf{acl}^{\mathsf{fr}}2.$ $\vdash \forall X, x\, \exists Y\, \forall y\; (y \in Y \leftrightarrow (y \in X \vee y = x))$,

$\mathsf{acl}^{\mathsf{fr}}3.$ $\vdash \forall X\, \exists x\; X \succ x$,

$\mathsf{acl}^{\mathsf{fr}}4.$ $\vdash \forall X, Y, z\; ((X \succ z \wedge Y \succ z) \to \forall w\; (w \in X \leftrightarrow w \in Y))$.

Remark 4.1 If one adds full comprehension to $\mathsf{acl}^{\mathsf{fr}}$, one can derive the Russell paradox as follows. Let R be a class such that, for all x, we have x is in R iff, for no X, $x \in X$ and $X \succ x$. Pick an r such that $R \succ r$. In case $r \in R$, we have $r \in R$ and $R \succ r$. So, $r \notin R$. A contradiction. So $r \notin R$. Suppose, for some X, we have $r \in X$ and $X \succ r$. Since $X \succ r$, we find that X is extensionally equal to R. Ergo $r \notin X$. A contradiction. So, for all X, not ($r \in X$ and $X \succ r$). This tells us that $r \in R$. A contradiction on no assumptions. ❏

We prove the following theorem.

Theorem 4.2 $\mathsf{acl}^{\mathsf{fr}} \equiv_{\mathsf{dir}} \mathsf{AS}$.

We will prove the theorem by showing that $\mathsf{acl}^{\mathsf{fr}} \equiv_{\mathsf{dir}} \mathsf{AS}^{\mathsf{s}}$. We directly interpret AS^{s} in $\mathsf{acl}^{\mathsf{fr}}$ via the following translation $\flat$.

- $\sigma_\flat \mathfrak{o} := \mathfrak{o}$,
- $\delta_\flat^{\mathfrak{o}} v_0 :\leftrightarrow v_0 = v_0$,
- $x E_\flat y :\leftrightarrow x = y$,
  $\mathsf{set}_\flat(x) :\leftrightarrow \exists X\; X \succ x$,
  $x \in_\flat y :\leftrightarrow \exists Y\; (x \in Y \wedge Y \succ y)$.

It is easily seen that indeed $\flat$ supports an interpretation of AS^{s} in $\mathsf{acl}^{\mathsf{fr}}$. Conversely, we have a direct interpretation of $\mathsf{acl}^{\mathsf{fr}}$ into AS^{s}, which is given by the translation $\sharp$.

- $\sigma_\sharp \mathfrak{c} := \sigma_\sharp \mathfrak{o} := \mathfrak{o}$,
- $\delta_\sharp^{\mathfrak{o}} v_0 :\leftrightarrow v_0 = v_0$, $\delta_\sharp^{\mathfrak{c}} v_0 :\leftrightarrow \mathsf{set}_0(v_0)$,
- $x E_\sharp^{\mathfrak{o}} y :\leftrightarrow x = y$, $x \in_\sharp y :\leftrightarrow x \in y$, $x \succ_\sharp y :\leftrightarrow (\mathsf{set}(x) \wedge x = y)$.

Under the weak assumptions we are considering, $\flat$ and $\sharp$ are in no sense each other inverses. This can be seen e.g. by noting that the Frege relation produced by $\sharp$ is functional. We can do a little bit better by adding the following axiom to AS^{s}.

- $\vdash \forall X, Y, z\; (\forall w\; (w \in X \leftrightarrow w \in Y) \to (X \succ z \to Y \succ z))$.

Clearly our axiom justifies considering extensional equality as an identity relation. Thus, we can definitionally extend $\mathsf{acl}^{\mathsf{fr}}$ with an identity relation on classes defined by extensional equality. We now can extend $\sharp$ by taking

$$xE^{\mathsf{c}}y :\leftrightarrow (x, y : \mathsf{set} \wedge \forall z\ (z \in x \leftrightarrow z \in y)).$$

One can show that the modified $\flat$ and $\sharp$ are each other's inverse in the sense that they witness bi-interpretability. See [Vis06], for the relevant notions.

4.2 Adjunctive Set Theory and Pairing

We consider a theory of non-surjective unordered pairing and a theory of non-surjective ordered pairing. Our theory of unordered pairing[13] has the same signature as AS. We define $\mathsf{PAIR}^{\mathsf{uno,ns}}$ by:

$\mathsf{PAIR}^{\mathsf{uno,ns}}1.\ \vdash \exists y\, \forall x\ x \notin y$,

$\mathsf{PAIR}^{\mathsf{uno,ns}}2.\ \vdash \forall x, y\, \exists z\, \forall u\ (u \in z \leftrightarrow (u = x \vee u = y))$.

Our theory of ordered pairs is a one-sorted theory that has, apart from identity, a single ternary predicate pair. (Alternatively, we could develop it with two partial projection functions π_0 and π_1.)

$\mathsf{PAIR}^{\mathsf{o,ns}}1.\ \vdash \exists z\, \forall x, y\ \neg\mathsf{pair}(x, y, z)$,

$\mathsf{PAIR}^{\mathsf{o,ns}}2.\ \vdash \forall x, y\, \exists z\ \mathsf{pair}(x, y, z)$,

$\mathsf{PAIR}^{\mathsf{o,ns}}3.\ \vdash \forall x, x', y, y', z\ ((\mathsf{pair}(x, y, z) \wedge \mathsf{pair}(x', y', z)) \to (x = x' \wedge y = y'))$.

We can directly interpret $\mathsf{PAIR}^{\mathsf{uno,ns}}$ in $\mathsf{PAIR}^{\mathsf{o,ns}}$, by translating $x \in y$ to the formula $\exists u\ (\mathsf{pair}(x, u, y) \vee \mathsf{pair}(u, x, y))$. We can directly interpret $\mathsf{PAIR}^{\mathsf{o,ns}}$ in $\mathsf{PAIR}^{\mathsf{uno,ns}}$ via Wiener-Kuratowski pairing. We translate $\mathsf{pair}(x, y, z)$ into:

$$\begin{aligned}&\exists u, v\ (\ \forall w\ (w \in z \leftrightarrow (w = u \vee w = v)) \wedge \\ &\forall w'\ (w' \in u \leftrightarrow w' = x) \wedge \forall w''\ (w'' \in v \leftrightarrow (w'' = x \vee w'' = y))).\end{aligned}$$

We consider $\mathsf{CART} := \mathsf{acl} \boxplus \mathsf{PAIR}^{\mathsf{o,ns}}$. We clearly have that $\mathsf{AS} \rhd_{\mathsf{dir}} \mathsf{CART}$. The desired translation uses the Wiener-Kuratowski formula given above. We will prove the converse.

Theorem 4.3 $\mathsf{CART} \rhd_{\mathsf{dir}} \mathsf{AS}$, *and, hence,* $\mathsf{CART} \equiv_{\mathsf{dir}} \mathsf{AS}$.

Proof

To give the heuristic, let's ignore for a moment the fact that pairing is not necessarily functional. The basic idea is to code e.g. the set consisting of a, b, c as $\langle\langle\langle 0, a\rangle, b\rangle, c\rangle$, where 0 is a non-pair. Now forget about functionality again. We define:

[13] John Burgess in [Bur05] calls the result of adding extensionality to this theory: UST.

- $\mathsf{dc}(Y) :\leftrightarrow \forall u, v, p\ ((\mathsf{pair}(u, v, p) \wedge p \in Y) \to u \in Y)$,
  (We will also write $Y{:}\mathsf{dc}$ for $\mathsf{dc}(Y)$.)

- $x \in y :\leftrightarrow \forall Y{:}\mathsf{dc}\ (y \in Y \to \exists w, q\ (\mathsf{pair}(w, x, q) \wedge q \in Y))$.

Consider any non-pair z. We clearly have $\{z\}{:}\mathsf{dc}$. If we would have $x \in z$, then, for some pair q, we would have that q is in the class $\{z\}$, quod non. So z is an empty set.

Consider any x and y. Pick any p with $\mathsf{pair}(y, x, p)$. We show that:

$$\forall u\ (u \in p \leftrightarrow (u \in y \vee u = x)).$$

We first treat the right-to-left direction. Suppose $u \in y$, $\mathsf{dc}(\mathrm{Y})$ and $p \in Y$. We find $y \in Y$, and, hence, for some w and q, $\mathsf{pair}(w, u, q)$ and $q \in Y$. So $u \in p$. Moreover, it is immediate that $x \in p$.

Conversely, suppose $u \in p$, $\mathsf{dc}(Y)$ and $y \in Y$. We have $\mathsf{dc}(Y \cup \{p\})$ and $p \in Y \cup \{p\}$. It follows that, for some w and q, we have $\mathsf{pair}(w, u, q)$ and $q \in Y \cup \{p\}$. If $q = p$, then $u = x$. If $q \neq p$, then $q \in Y$ and, thus, $u \in y$. ❑

The above theorem illustrates that the 'direct' sum $\boxplus$ makes the summands interact in non-trivial ways. In this, it contrasts with the ordinary disjoint sum $\oplus$. E.g., sequential theories like AS are connected or join-irreducible in the degrees of interpretability. See [Pud83] and [Ste89].

Corollary 4.4 *The theory* acl *is not directly interpretable in* $\mathsf{identity}$*, the theory of pure identity.*

Proof

Suppose $\mathsf{identity} \rhd_{\mathsf{dir}} \mathsf{acl}$. Then:

$$\begin{array}{rcl} \mathsf{PAIR}^{\mathsf{o,ns}} & = & \mathsf{identity} \boxplus \mathsf{PAIR}^{\mathsf{o,ns}} \\ & \rhd_{\mathsf{dir}} & \mathsf{acl} \boxplus \mathsf{PAIR}^{\mathsf{o,ns}} \\ & = & \mathsf{CART} \\ & \equiv_{\mathsf{dir}} & \mathsf{AS} \end{array}$$

However, $\mathsf{PAIR}^{\mathsf{o,ns}}$ has a decidable extension (see e.g., [Ten] or [CR01]) and AS is essentially undecidable. Quod impossibile. ❑

The theory CART is an ideal starting point to develop cardinal arithmetic in the traditional way. The Cartesian product will be an appropriate class of pairs. Note however that it is still a lot of work to get things done here. First, we have to live with the fact that our products are not unique, since we do not have functional pairing. Secondly, to get our product total, we need associativity of the product modulo isomorphism. Thus, we have to perform Solovay trickery to get the product associative.

We will follow another strategy. We will develop cardinal arithmetic using the universal properties of sums and products that are employed in category theory. As a bonus it turns out that this plan can be executed in a theory COPY that is strictly weaker than CART in the sense of direct interpretability. The theory COPY will be constructed using Adjunctive Relation Theory, which we present in the next section.

5 Adjunctive Relation Theory

We define adjunctive relation theory, arel, as follows. The theory arel is two-sorted, with a sort $\mathfrak{o}$ of objects and a sort $\mathfrak{r}$ of binary relations. We have a ternary application predicate app of type $\mathfrak{roo}$. We write 'Rxy' or '$(x,y) \in R$' for: $\mathsf{app}(R,x,y)$.

$\mathsf{arel}1.$ $\vdash \exists R\, \forall x, y\, \neg Rxy$,

$\mathsf{arel}2.$ $\vdash \forall R, x, y\, \exists S\, \forall u, v\, (Suv \leftrightarrow (Ruv \vee (u = x \wedge v = y)))$.

Adjunctive Relation Theory is the theory employed by John Burgess in [Bur05], section 2.2, to interpret Q. Specifically, let SUCC be the theory of one single successor operation and zero, given by the first two axioms of Q. Burgess shows how to interpret Q in $\mathsf{SUCC}^+ := \mathsf{arel} \boxplus \mathsf{SUCC}$.

We determine the relationships between arel, acl and $\mathsf{identity}$.

Theorem 5.1 *We have:*

a. $\mathsf{identity} \rhd \mathsf{arel}$,

b. $\mathsf{arel} \rhd_{\mathsf{dir}} \mathsf{acl}$ *and* $\mathsf{acl} \rhd_{\mathsf{dir}} \mathsf{identity}$,

c. $\mathsf{acl} \not\rhd_{\mathsf{dir}} \mathsf{arel}$ *and* $\mathsf{identity} \not\rhd_{\mathsf{dir}} \mathsf{acl}$.

Thus, there is a strictly ascending sequence of three in the degrees of direct interpretability that is contained in the minimal degree of interpretability.

Proof

We have (a), because arel has a finite model. (b) is easy. The second claim of (c) is Corollary 4.4. We prove the first claim of (c). Suppose $\mathsf{acl} \rhd_{\mathsf{dir}} \mathsf{arel}$. Then,

$$\begin{array}{rcl} \mathsf{acl} \boxplus \mathsf{SUCC} & \rhd_{\mathsf{dir}} & \mathsf{arel} \boxplus \mathsf{SUCC} \\ & = & \mathsf{SUCC}^+ \\ & \rhd & \mathsf{Q} \end{array}$$

It follows that $\mathsf{acl} \boxplus \mathsf{SUCC}$ is essentially undecidable. On the other hand, $\mathsf{acl} \boxplus \mathsf{SUCC}$ is contained in the (true) monadic second order theory of one successor. This theory is decidable. A contradiction.

Alternatively, $\mathsf{acl} \boxplus \mathsf{SUCC}$ is contained in the weak (true) monadic second order theory of one successor. Here the second order variables range over finite sets. This theory is decidable. Again we have our contradiction.

We may conclude that arel is not directly interpretable in acl.

See [BE59], [Büc60], [Elg61], [ER66], for the basics on weak and strong successor theories. ❑

5.1 Boolean Operations on Relations

We can repeat the development in Section 3, to strengthen our theory with identity between relations and total operations of disjunction and intersection on relations. Moreover, we can get, via local direct interpretability, for any theory U, an appropriate version of $\Delta_0^{\subseteq}(U)$-Aussonderung in $\mathsf{arel} \boxplus U$. Note that this version of Aussonderung guarantees the existence of relations of the form $\{(x,y) \in R \mid A(x,y,\vec{z},\vec{S})\}$, where all relation quantifiers in A are subrelation bounded. We call the theory $\mathsf{arel} \boxplus U$, plus the axiom *unions of relations exist*, plus $\Delta_0^{\subseteq}(U)$-Aussonderung, $\mathsf{arel}^+(U)$. If U is the theory of identity, we will simply use arel^+.

5.2 Relational Algebra

We enrich $\mathsf{arel}^+(U)$ to a theory with several convenient operations on relations. We define:

- $x\ \delta(R)\ y :\leftrightarrow x = y \wedge \exists z\ xRz$,
- $x\ \rho(R)\ y :\leftrightarrow x = y \wedge \exists z\ zRx$,
- $x\ \gamma(R)\ y :\leftrightarrow yRx$,
- $x\ (R \circ S)\ y :\leftrightarrow \exists z\ (xSz \wedge zRy)$.

First, we take rel^0 as the class of all relations R such that $\gamma(R)$ exists.

Theorem 5.2 ($\mathsf{arel}^+(U)$) rel^0 *is closed under empty relation, singleton relations, union, and γ and is downwards closed w.r.t. $\subseteq$.*

Proof

Closure under the empty relations and singleton relations is easy. Clearly γ commutes with unions, and hence rel^0 is closed under union. Closure under γ follows from the fact that $\gamma\gamma(R) = R$. ❑

We relativise our relations to rel^0. In the resulting theory, say $\mathsf{arel}_1^+(U)$, γ is total. We proceed in $\mathsf{arel}_1^+(U)$. Let rel^1 is the class of all relations R such that:

1. $\delta(R)$ exists,
2. $\rho(R)$ exists,
3. for all S, we have $S \circ R$ and $R \circ S$ exist,
4. there is an $R' \subseteq R$, R' is functional and $\delta(R') = \delta(R)$,
5. there is an $R' \subseteq R$, R' is injective and $\rho(R') = \rho(R')$,

We have:

Theorem 5.3 ($\mathsf{arel}_1^+(U)$) *The predicate* rel^1 *is closed under empty relation, singleton relations, union, and is downwards closed w.r.t.* $\subseteq$. *Moreover, it is closed under* δ, ρ, γ *and composition*

Proof

The case of the empty relation and of singleton relations is easy.

We treat *downward closure.* Suppose $R^* \subseteq R$. The cases of δ, ρ and composition are by $\Delta_0^{\subseteq}$-Aussonderung. E.g., $\delta(R^*) = \{(x, y) \in \delta(R) \mid \exists z\ xR^*z\}$. If R' is the functional subrelation guaranteed for R, then $\{(x, y) \in R' \mid \exists z\ xR^*z\}$ is the one we are looking for in the case of R^*. The case of the injective subrelation is similar.

We show that rel^1 *is closed under union.* Suppose R_0 and R_1 are in rel^1. Clearly, δ and ρ, and composition commute with union. So $\delta(R_0 \cup R_1)$, is in rel^1, and similarly for ρ and composition. Let $R_2 := \{(x, y) \in R_1 \mid \neg\,\exists z\ R_0xz\}$. Then, R_2 is in rel^1. Let R_0', R_2' be the promised functional subrelations for respectively R_0, and R_2. Then $R_0' \cup R_2'$ is a functional subrelation for $R_0 \cup R_2$, which is identical to $R_0 \cup R_1$. The case of the injective subrelation is similar.

We show that rel^1 *is closed under* δ *and* ρ. Suppose R is in rel^1. Then $\delta(R)$ exists. Clearly, $\delta\delta(R) = \delta(R)$, $\rho\delta(R) = \delta(R)$ and $\gamma\delta(R) = \delta(R)$, so $\delta\delta(R)$, $\rho\delta(R)$ and $\gamma\delta(R)$ exist. Consider any S. Clearly $S \circ \delta(R)$ and $\delta(R) \circ S$ are subrelations of S. Hence they exist. Since $\delta(\rho)$ is both functional and injective, we can take it as its own desired subrelation. The case of ρ is similar.

We prove that rel^1 *is closed under* γ. Suppose R is in rel^1. We show that $\gamma(R)$ is in rel^1. Clearly, $\delta\gamma(R) = \rho(R)$, $\rho\gamma(R) = \delta(R)$ and $\gamma\gamma(R) = R$. So, $\delta\gamma(R)$, $\rho\gamma(R)$ and $\gamma\gamma(R)$ exist. Consider any S. We have $\gamma(R) \circ S = \gamma(\gamma(S) \circ R)$, and, hence, $\gamma(R) \circ S$ exists. Similarly for $S \circ \gamma(R)$. Finally, let R' be an injective subrelation of R with $\rho(R') = \rho(R)$. Then, $\gamma(R')$ is a functional subrelation of $\gamma(R)$ with $\delta\gamma(R') = \delta\gamma(R)$. Similarly, for the case of the injective subrelation. ❑

Relativising to rel^1 and adding the new operations to the signature, we obtain the theory $\mathsf{arel}_2^+(U)$ which is $\mathsf{arel}_0^+(U)$ plus the defining axioms and the totality of δ, ρ, γ, $\circ$, the principle that every relation has a functional subrelation with the same domain and the principle that every relation has an injective subrelation with the same range.

5.3 Adding Classes

We proceed to add classes to our theory $\mathsf{arel}_2^+(U)$. We extend the signature with a new sort $\mathfrak{c}$, identity for the new sort, a relation symbol ϵ of sort $\mathfrak{oc}$, a function symbol diag of sort $\mathfrak{cr}$ and function symbols δ^+ and ρ^+ of sort $\mathfrak{rc}$. We define the domain for $\mathfrak{c}$ to be class, where $\mathsf{class}(R) :\leftrightarrow \forall x, y\ (xRy \to x = y)$. We take, for R in class, $\mathsf{diag}(R) := R$, and $\delta^+(R) := \delta(R)$, $\rho^+(R) := \rho(R)$.

The axioms for $\mathsf{arel}_3^+(U)$ are the axioms for $\mathsf{arel}_2^+(U)$, plus the functionality axioms for the new operations and:

- $\vdash \mathsf{diag}(X)(x, y) \leftrightarrow (x \in X \wedge x = y)$,
- $\vdash X = Y \leftrightarrow \mathsf{diag}(X) = \mathsf{diag}(Y)$,
- $\vdash \mathsf{diag}(\delta^+(R)) = \delta(R)$,
- $\vdash \mathsf{diag}(\rho^+(R)) = \rho(R)$.

Note that we can derive union and $\Delta_0^{\subseteq}(U)$-Aussonderung for classes from these axioms.

We definitionally expand $\mathsf{arel}_3^+(U)$ with three operations:

- $R \restriction X := \{(x, y) \in R \mid x \in X\}$.
- $R \upharpoonleft Y := \{(x, y) \in R \mid y \in Y\}$.
- $R \otimes S := R \cup (S \restriction (\delta(S) \setminus \delta(R)))$.

These operation are total by $\Delta_0^{\subseteq}(U)$-Aussonderung. The resulting theory is $\mathsf{arel}_4^+(U)$. Note that our operations preserve functionality.

5.4 Building a Category

We proceed to build a category in $\mathsf{arel}_4^+(U)$. We add a sort of morphisms. We define morphisms $f, g, h, \ldots$ via a three dimensional interpretation. It's domain will be given by the formula morph. We say $\mathsf{morph}(X, F, Y)$ iff X and Y are classes and F is a functional relation with $\delta^+(F) = X$ and $\rho^+(F) \subseteq Y$. We let the variable f correspond to the triple (X, F, Y), and g to (W, G, Z) and h to (U, H, V). We add various notations. Let's call our translation $\digamma$. We translate all our existing predicats like $\in$ to themselves.

- $f : A \to B$ for: $X = A$ and $Y = B$. Officially: $(f : A \to B)_{\digamma}$ iff $X = A$ and $Y = B$.

- $f = g$ for $X = W$ and $Y = Z$ and $F = G$.
- $\mathsf{dom}(f) = A$ for: $X = A$.
- $\mathsf{cod}(f) = B$ for: $Y = B$.
- $\mathsf{carr}(f) = J$ for: $F = J$.
- $fx = y$ for: Fxy.
- id_A is the triple $(A, \mathsf{diag}(A), A)$. Officially it is defined as: $(\mathsf{id}_A(f))_F$ iff $X = A$, $Y = A$ and, $F = \mathsf{diag}(A)$. Note that id_A is a constant: there is a unique f satisfying $(\mathsf{id}_A(f))_F$.
- $\mathsf{co}_{A,a}$ is a partial constant denoting the function from A to $\{a\}$ which sends every $a' \in A$ to a. Officially it is defined as: $(\mathsf{co}_{A,a}(f))_F$ iff $X = A$, $Y = \{a\}$ and, for all a' in A, $Fa'a$.
- ini_A is the unique function from $\emptyset$ to A, so $(\mathsf{ini}_A(f))_F$ iff $X = \emptyset$ and $F = \emptyset$ and $Y = A$.
- $((f \restriction A)(g))_F$ iff $W = (X \cap A)$ and $Z = Y$ and $G = F \restriction A$.
- $((f \upharpoonleft B)(g))_F$ iff $W = \{x{\in}X \mid f(x) \in B\}$ and $Z = (Y{\cap}B)$ and $G = F \upharpoonleft B$.
- $f \circledvee g$ is given by: $((f \circledvee g)(h))_F$ iff $U = (X \cup W)$ and $V = (Y \cup Z)$ and $H = F \circledvee G$.
- $f \circ g$ is given by $((f \circ g)(h))_F$ iff $U = W$ and $V = Y$ and $H = F \circ G$.

Note that all our definitions can be rewritten to $\Delta_0^{\subseteq}(U)$-formulas.

We work in $\mathsf{CAT}_0(U) := F^{-1}(\mathsf{arel}_4^{+}(U))$. Since via dom, cod, application and carr, we have access to the underlying triples, we can switch from a morphism f to an underlying presentation. Note that we have, for X, Y and F, such that $X = \delta(F)$ and $\rho(F) \subseteq Y$, that there is a unique f such that $\mathsf{dom}(f) = X$, $\mathsf{cod}(f) = Y$ and $\mathsf{carr}(f) = F$. Thus, we may speak about the morphism f given by (X, F, Y).

The strategy we will follow to gain more properties for our category is to restrict the classes to some totality, say $\mathcal{X}$, that is closed under union and downward closed under $\subseteq$. We restrict our relations to those relations that have domain and range in $\mathcal{X}$, and we similarly restrict our morphisms. It is easily seen that the new relations are closed under union, taking subrelations and under the operations of Subsection 5.2. Similarly, the new morphisms will be closed under the operations specified above. It follows that the theory obtained by redefining our sorts using these predicates, still satisfies all the desired properties.

Let class^0 consist of all the X such that, for any y, the unique function G from X to $\{y\}$ exists. We have the following easy theorem.

Theorem 5.4 ($\mathsf{CAT}_0(U)$) class^0 *is closed under empty class, singletons, union and is downward closed under the subset relation.*

We form our next theory $\mathsf{CAT}_1(U)$ by restricting our relations and morphisms to those with domain and codomain from class^0 as described above. Note that $\mathsf{co}_{A,a}$ is inhabited in $\mathsf{CAT}_1(U)$.

We can prove many properties familiar form the category of sets in $\mathsf{CAT}_1(U)$. Here are a few of these.

- f is an injection iff f is a monomorphism iff f is a split monomorphism.
- f is a surjection iff f is an epimorphism iff f is a split epimorphism.
- f is a bijection iff f is an isomorhism.
- We have an initial object and we have end objects.
- We have all equalizers.

5.5 The Embedding Ordering on Classes

The embedding ordering between classes will be useful in the development of the product, following an idea due to Mycielski, Pudlák and Stern (see Appendix III of [MPS90]). In this subsection we will prove an important fact about this ordering.

We work in $\mathsf{CAT}_1(U)$. We define $X \preceq Y$ iff there is a injection $f : X \to Y$.

Theorem 5.5 ($\mathsf{CAT}_1(U)$) *a. $\preceq$ is a partial preordering.*

b. Suppose that $X \subseteq Y$. Then, $X \preceq U$.

Proof

It is easy to see that $\preceq$ is a partial preordering. Let $\mathsf{e}_{XY} := (\mathsf{id}_X \circledcirc \mathsf{ini}_Y)$. Then $\mathsf{e}_{X,Y}$ is an injection, witnessing $X \preceq Y$. ❑

We have the following.

Theorem 5.6 ($\mathsf{CAT}_1(U)$) *Suppose a virtual class $\mathcal{J}$ of classes is closed under empty class, singletons and union. Let $\mathcal{J}^{\mathsf{pr}}$ be the virtual class of all X such that, for all Y, if $Y \preceq X$, then Y is in $\mathcal{J}$. Then, $\mathcal{J}^{\mathsf{pr}}$ is closed under empty class, singletons and union. Moreover, $\mathcal{J}^{\mathsf{pr}}$ is downwards closed under $\preceq$, and, hence under $\subseteq$.*

Proof

Let $\mathcal{J}$ be as stipulated in the theorem. It is easy to see that $\mathcal{J}^{\mathsf{pr}}$ is closed under empty class and singletons.

Suppose X_0 and X_1 are in $\mathcal{J}^{\mathsf{pr}}$. Let $X := X_0 \cup X_1$. Suppose $Y \preceq X$. Let $f : Y \to X$ be an injection. Take $f_i := f \upharpoonright X_i$ and let $Y_i := \mathsf{dom}(f_i)$. then,

$f_i : Y_i \to X_i$ is an injection. So, $Y_i \preceq X_i$. It follows that the Y_i are in $\mathcal{J}$. It is easily seen that $Y = Y_0 \cup Y_1$. Hence, Y is in $\mathcal{J}$.

Suppose $Z \preceq Y \preceq X$ and X is in $\mathcal{J}^{\mathsf{pr}}$. Let $g : Z \to Y$ and $f : Y \to X$ be injective. Then, $f \circ g : Z \to X$ is injective. Hence $Z \preceq X$, and, thus, $Z \in \mathcal{J}$.

Suppose $Y \subseteq X$ and X is in $\mathcal{J}^{\mathsf{pr}}$. Then, $Y \preceq X$ and, hence Y is in $\mathcal{J}^{\mathsf{pr}}$. ❑

6 Constructing Sum and Product

It is easily seen that we cannot expect to have all sums in $\mathsf{CAT}_1(U)$, for all U. E.g., if we take U to be $\mathsf{identity}$, we see that our theory has finite models. Thus, we need a suitable base theory to get sums. Our choice is the theory COPY. In this section we construct sum and product in COPY.

6.1 Introducing COPY

The theory TJI of two total, jointly injective relations is one-sorted and has to binary predicates P_0 and P_1. The theory is axiomatized as follows.

TJI1. $\vdash \forall x \, \exists y \; P_i(x, y) \;\; (i = 0, 1)$,

TJI2. $\vdash (P_i(x, y) \wedge P_i(x', y)) \to x = x' \;\; (i = 0, 1)$,

TJI3. $\vdash \neg \, (P_0(x, y) \wedge P_1(x', y))$.

Now consider $\mathsf{COPY} := \mathsf{arel} \boxplus \mathsf{TJI}$.

Theorem 6.1 CART *directly interprets* COPY*, but* COPY *does not directly interpret* CART.

Proof

We specify a direct interpretation of COPY in CART. We interpret objects as objects and relations as classes. We translate:

- Rxy to $\exists p \, (\mathsf{pair}(x, y, p) \wedge p \in R)$.
- $P_0(x, y)$ to $\exists z \, (\forall u, v \, \neg \, \mathsf{pair}(u, v, z) \wedge \mathsf{pair}(x, z, y))$,
- $P_1(x, y)$ to $\exists z \, (\mathsf{pair}(x, x, z) \wedge \mathsf{pair}(x, z, y))$.

To see that COPY does not directly interpret CART, consider the following model $\mathcal{M}$ of COPY. The object domain of $\mathcal{M}$ is $\omega \times \omega$. We set:

- $P_i^{\mathcal{M}}(\langle m, n \rangle, \langle p, q \rangle) :\Leftrightarrow m = p$ and $q = 2n + i \quad (i = 0, 1)$.

The relations of $\mathcal{M}$ are the finite relations on $\omega \times \omega$. Suppose that some formula A represents pair in $\mathcal{M}$ in finitely many object and relation parameters. The set X of first components of the object parameters and of the elements of the domains and ranges of the relation parameters is finite. Suppose m, n, k, k' are pairwise disjoint numbers, not in X. For some $\langle p, q \rangle$, we have $\mathcal{M} \models A(\langle m, 0 \rangle, \langle n, 0 \rangle, \langle p, q \rangle)$. (We suppress the parameters in A.) Let σ be the function that interchanges m and k on ω. Let $\sigma^\star \langle i, j \rangle := \langle \sigma i, j \rangle$. Clearly, $\sigma^\star$ lifts to an automorphism of $\mathcal{M}$ that leaves the parameters in place. We have $\mathcal{M} \models A(\langle k, 0 \rangle, \langle n, 0 \rangle, \langle \sigma p, q \rangle)$. It follows that $p = m$ or $p = k$. Since we could have chosen k' for k, we may conclude that $p = m$. By similar reasoning, we find $p = n$. A contradiction. ❑

Note that the model used in the proof is a weak standard model, since the class variables range over all finite sets of objects. The argument would have worked as well if we had taken all sets as the classes of the model and, thus, had considered a strong standard model. So there are both weak and strong standard models of COPY, such that pairing of objects cannot be defined in these models.

Open Question 6.2 What are the precise relations in the preorder of direct interpretability of the theories COPY and $\mathsf{arel} \boxplus \mathsf{SUCC}$? ❑

6.2 Building Isomorphic Copies of Classes

Our theory COPY is $\mathsf{arel} \boxplus \mathsf{TJI}$. Using the materials of Subsection 5.4, we build a theory $\mathsf{CAT}_0^\star := \mathsf{CAT}_1(\mathsf{TJI})$, that is locally directly interpretable in COPY. We work in $\mathsf{CAT}_0^\star$.

We define class^0 to be the totality of all X, such that for $i = 0, 1$, there is a function F_i with $\delta^+(F_i) = X$, such that, for all $x \in X$, $P_i(x, F_i x)$.

Theorem 6.3 $(\mathsf{CAT}_0^\star)$ class^0 *is closed under empty class, singletons, union and is downward closed under the subset relation.*

Proof

Closure under empty class and singletons is trivial.

We prove closure under subsets. Suppose X is in class^0 and $X' \subseteq X$. Let F_i be the functions promised for X in the first clause. We may take $F_i' := F_i \restriction X'$.

We prove closure under unions. Suppose X_0 and X_1 are in class^0. Let $X := X_0 \cup X_1$. Let F_{ji} be the functions promised for X_j. We may take as the desired functions for X the functions $F_i := F_{0i} \circledvee F_{1i}$. ❑

Consider X. Suppose there are functions F_i with $\delta^+(F_i) = X$, such that, for all $x \in X$, $P_i(x, F_i x)$. Let $Y_i := \rho(F_i)$. since P_i is injective, it follows that F_i, is injective. Thus, the morphism f_i given by (X, F_i, Y_i) is a bijection. It follows that f_i is an isomorphism, since f_i has an inverse given by $(Y_i, \gamma(F_i), X)$. Note that the Y_i are disjoint, by the joint injectivitiy of the P_i.

We define class^1 to be $(\mathsf{class}^0)^{\mathsf{pr}}$, i.e., the class of all X, such that, for all $Y \preceq X$, we have Y is in class^0. By Theorem 5.6, class^1 will be closed under union and downward closed under $\preceq$. We restrict our classes to class^1, we restrict our relations to the relations with domain and range in class^1 and we similarly restrict our morphisms. Thus, we obtain $\mathsf{CAT}_1^\star$. We preserve all the desirable properties we previously acquired.

Note that, for X in class^1, it follows that the isomorphic copies Y_i are also in class^1. Thus, $\mathsf{CAT}_1^\star$ proves that, for any X, there are mutually disjoint isomorphic copies Y_i. What is more, for any X_0, X_1, we have an isomorphic copy Y_{00} of X_0 and an isomorphic copy Y_{11} of X_1 such that Y_{00} and Y_{11} are mutually disjoint.

6.3 The Sum

We show that sums exist in $\mathsf{CAT}_1^\star$. We build a sum $X_0 + X_1$ of X_0 and X_1 as follows. Let X_0' and X_1' be disjoint isomorphic copies of X_0, resp. X_1. We claim that $X_0' \cup X_1'$ is a sum of X_0 and X_1 with as in-arrows $g_i \circledvee \mathsf{ini}_{X_{1-i}}$, where $g_i : X_i \to X_i'$ is the standard isomorphism with inverse h_i.

Suppose $f_i : X_i \to Y$. Let $g_i := (f_i \circ h_i) : X_i' \to Y$. The unique arrow finishing the sum diagram is $g_0 \circledvee g_1$. The verification is the usual reasoning.

Note that in models of our theory we will have a functor +. However, our theory cannot 'see' this, since we do not have enough Choice in the theory to select unique representatives.

6.4 The Product

We work again in $\mathsf{CAT}_1^\star$. Let class^2 be the virtual class of the X such that, for all Y a cartesian product $X \times Y$ exists. Here the product is defined by its usual universal property.

Theorem 6.4 *The virtual class* class^2 *is closed under empty class, singletons, sum, product, and is downwards closed w.r.t.* $\preceq$. *It follows that* class^2 *is closed under unions.*

Proof

The case of the empty class is easy. For *the case of singletons* we can take $\{x\} \otimes Y$ to be Y, with $\pi_0^{\{x\},Y} := \mathsf{co}_{Y,x}$ and $\pi_1^{\{x\},Y} := \mathsf{id}_Y$.

We prove closure under sum. Consider three classes X_0, X_1 and Y. $X_0 + X_1$ is isomorphic to the union of two disjoint isomorphic copies X_0' and X_1' of, respectively, X_0 and X_1. Since the product is determined modulo isomorphism anyway, it is sufficient to prove our theorem for this disjoint union. So, we assume that X_0 and X_1 are disjoint and that $X_0 + X_1$ is their union.

Suppose $X_i \times Y$ exists with witnessing projection functions π^i_j. We show that any sum $(X_0 \times Y) + (X_1 \times Y)$ is a product $(X_0 + X_1) \times Y$ with witnessing projection functions $\pi^*_0 := \pi^0_0 + \pi^1_1$ and $\pi^*_1 := [\pi^1_0, \pi^1_1]$. Here $[\cdot, \cdot]$ is given by the sum diagram.

$$\begin{tikzcd}
X_0 \times Y \arrow[r, "\mathsf{in}_0"] \arrow[dr, "\pi^1_0"'] & (X_0 \times Y) + (X_1 \times Y) \arrow[d, dotted, "{[\pi^1_0, \pi^1_1]}"] & X_1 \times Y \arrow[l, "\mathsf{in}_1"'] \arrow[dl, "\pi^1_1"] \\
& Y &
\end{tikzcd}$$

Suppose we have $f : W \to X_0 + X_1$ and $g : W \to Y$. Let $f_i := f \upharpoonright X_i$ and $W_i := \mathsf{dom}(g_i)$. Clearly the W_i form a partition of W. Let $g_i := g \upharpoonright W_i$. We find: $f = f_0 + f_1$ and $g = [g_0, g_1]$. Clearly we have $(f_i, g_i) : W \to X_i \times Y$. Here $(\cdot, \cdot)$ is given by the product diagram.

$$\begin{tikzcd}
& W_i \arrow[dl, "f_i"'] \arrow[d, dotted, "{(f_i, g_i)}"] \arrow[dr, "g_i"] & \\
X_i & X_i \times Y \arrow[l, "\pi^i_0"] \arrow[r, "\pi^i_1"'] & Y
\end{tikzcd}$$

Putting everything together, we obtain the following diagram.

$$\begin{tikzcd}
& W_0 + W_1 \arrow[dl, "f_0 + f_1"'] \arrow[d, dotted, "{(f_0, g_0) + (f_1, g_1)}"] \arrow[dr, "{[g_0, g_1]}"] & \\
X_0 + X_1 & (X_0 \times Y) + (X_1 \times Y) \arrow[l, "\pi^0_0 + \pi^1_0"] \arrow[r, "{[\pi^0_1, \pi^1_1]}"'] & Y
\end{tikzcd}$$

It is easily seen that the diagram commutes. Suppose h also finishes the diagram. We have: $h : (W_0 + W_1) \to ((X_0 \times Y) + (X_1 \times Y)$. Since, we may work modulo isomorphism, we may assume $(X_0 \times Y)$ and $(X_1 \times Y)$ to be disjoint and $(X_0 \times Y) + (X_1 \times Y)$ to be the union of $(X_0 \times Y)$ and $(X_1 \times Y)$. Let $h_i := h \upharpoonright W_i$. Clearly the range of h_i is $X_i \times Y$. So, $h = h_0 + h_1$. We may conclude that $\pi^i_0 \circ h_i = f_i$ and $\pi^i_1 \circ h_i = g_i$. So $h_i = (f_i, g_i)$ and, thus, $h = (f_0, g_0) + (f_1, g_1)$.

We show that class^2 *is closed under products.* Suppose X_0 and X_1 are in class^2. Consider any Y. Since X_1 is in class^2, we find that $X_1 \times Y$ exists. Since, X_0

is in class^2, we find that $X_0 \times (X_1 \times Y)$ exists. Similarly, $X_0 \times X_1$ exists. By familiar arguments, $X_0 \times (X_1 \times Y)$ is also a product of the form $(X_0 \times X_1) \times Y$ with projections $\pi_0^{X_0 \times X_1, Y} = \mathsf{id}_{X_0} \times \pi_0^{X_1,Y}$ and $\pi_1^{X_0 \times X_1, Y} = \pi_1^{X_1,Y} \circ \pi_1^{X_0, X_1 \times Y}$.

We show that class^2 *is downward closed w.r.t.* $\preceq$. Suppose $f : W \to X$ is an injection and X is in class^2. Let $X' := f[W] := \{x{\in}X \mid \exists w{\in}W\ f(w) = x\}$. By $\Delta_0^{\subseteq}$-Aussonderung, X' exists. Clearly, $f' := f \upharpoonright X'$ is a bijection. Hence, f' is an isomorphism. Since class^2 is evidently closed under isomorphism, it is sufficient to show that X' is in class^2.

Consider any Y. We take $X' \times Y := Z := \{z{\in}(X \times Y) \mid \pi_0^{X,Y} \in X'\}$, and $\pi_i^{X',Y} := \pi_i^{X,Y} \upharpoonright Z$. Let $f' : U \to X'$ and $g : U \to Y$. We take $f := f' \circledcirc \mathsf{ini}_X$ and $(f', g) := (f, g) \upharpoonright Z$. It is easy to see that (f', g) makes the product diagram for X' and Y commute and that any other such morphism is equal to (f', g).

Finally, we show closure under unions. Consider any X_0 and X_1 in class^2. We have $(X_1 \setminus X_0) \subseteq X_0$, hence, $(X_1 \setminus X_0) \preceq X_0$. It follows that $(X_1 \setminus X_0)$ is in class^2. The union of X_0 and $(X_1 \setminus X_0)$ is a sum, and, thus, it is in class^2. ❑

We restrict our classes to class^2 and restrict our relations and morphisms accordingly, thus obtaining a theory $\mathsf{CAT}_2^\star$, which is closed under products. Note that, since equalizers trivially exist, we find that according to $\mathsf{CAT}_2^\star$ all finite limits exist. Note also that the proof of Theorem 6.4 gives us the distributivity of sum over product.

Open Question 6.5 I. Can we improve upon our construction to get also all coequalizers (and, hence, all finite colimits)?

II. Can we develop the theory of exponentiation as a partial function in a similar style?

III. Can we improve our interpretation to get the theorem of Cantor-Schröder-Bernstein?

❑

Remark 6.6 Note that our interpretation of sum and product reduces to the identity interpretation in standard models of both sorts. Thus, our definition gives us all cardinals lesser than or equal to the cardinality of the domain for any model of COPY, where the class variables range over all subsets of the domain, and it gives us all finite cardinals for any model of COPY, where the class variables range over all finite subsets of the domain. ❑

7 Interpretation of Robinson's Arithmetic

We have shown how to locally directly interpret in COPY a theory $\mathsf{CAT}_2^\star$ in which we have a category of functions and classes. We now build the following interpretation of Q in $\mathsf{CAT}^{\star}_2$.[14] The domain of the interpretation is formed by our classes. Identity is interpreted by isomorphism in our category. Successor is the sum of the given class and any singleton, Addition is sum and Multiplication is product. In short: we interpret arithmetic as a theory of cardinals. We remind the reader of the axioms of Q.

Q1. $\vdash \mathsf{S}x = \mathsf{S}y \to x = y$,

Q2. $\vdash 0 \neq \mathsf{S}x$,

Q3. $\vdash x = 0 \vee \exists y\, x = \mathsf{S}y$,

Q4. $\vdash x + 0 = x$,

Q5. $\vdash x + \mathsf{S}y = \mathsf{S}(x + y)$,

Q6. $\vdash x \times 0 = 0$,

Q7. $\vdash x \times \mathsf{S}y = x \times y + x$.

We verify Q1. Suppose $X + \{x\}$ is isomorphic to $Y + \{y\}$, say via f. Without loss of generality, we may assume that X and $\{x\}$ are disjoint and that sum is union. Similarly, for Y and $\{y\}$. In case $fx = y$, we find that $f \restriction X$ is an isomorphism between X and Y. If not, there are $x_0 \in X$, $y_0 \in Y$ such that $fx_0 = y$ and $fx = y_0$. Let g be the unique morphism such that $g : \{x_0\} \to \{y_0\}$. It is easily verified that $(f \restriction (X \setminus \{x_0\})) \circledcirc g$ is an isomorphism between X and Y. The other axioms are even easier.

Note that our interpretation yields (at least) the following extra principles: the associativity and commutativity of plus and times and the distributivity of times over plus.

If we interpret $x \leq y$ as $x \preceq y$, we also get the principle:

Q8. $\vdash x \leq y \leftrightarrow \exists z\, z + x = y$.

Of course, we can also get this for free by defining $x \leq y$ as $\exists z\, z + x = y$.

Remark 7.1 Some philosophical logicians think that a theory should satisfy certain minimal conditions to qualify as an arithmetic or a set theory. I never completely understood how one would want to make philosophical judgments of this kind. Our construction suggest one possible answer for the case of Q. To view Q as a theory of cardinals one needs a development in a theory COPY of cardinals that give us precisely Q. However, we see that any plausible development of cardinals will also give us associativity and commutativity of plus and times and distributivity of times over plus. So, in a way, Q is too weak. ❑

Open Question 7.2 Precisely which principles are validated by our interpretation of Q? ❑

[14] Note that this interpretation is not direct.

Open Question 7.3 Our development of $\mathsf{CAT}^{\star}_{2}$ switched back and forth from categorical reasoning to reasoning about the implementation as a matter of course. Inspecting e.g. the subproof of closure under sums in the proof of Theorem 6.4, we see that it can be rewritten to a proof wholly in terms of the category by noting first some natural $\mathsf{CAT}^{\star}_{1}$-verifiable properties of the category. In other words, we can put all reasoning that calls upon the implementation at the beginning of the argument.

This raises the question, which properties we could ask of our category such that (1) we can directly interpret such a category easily in COPY and (2) the verification of the interpretation of Q is wholly categorical. It seems that the main problem here is the injectivity of successor: this appears to be deeply unnatural from the point of view of category theory. ❑

References

[BE59] J.R. Büchi and C.C. Elgot. Decision problems of of weak second order arithmetics and finite automata, part I (abstact). *American Mathematical Society Notices*, 5:834, 1959.

[Büc60] J.R. Büchi. Weak second order arithmetic and finite automata. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 6:66–92, 1960.

[Bur05] John Burgess. *Fixing Frege.* Princeton Monographs in Philosophy. Princeton University Press, Princeton, 2005.

[CH70] G.E. Collins and J.D. Halpern. On the interpretability of Arithmetic in Set Theory. *The Notre Dame Journal of Formal Logic*, 11:477–483, 1970.

[CR01] P. Cégielski and D. Richard. Decidability of the natural integers equipped with the Cantor pairing function and successor. *Theoretical Computer Science*, 257, 1–2:51–77, 2001.

[Elg61] C.C. Elgot. Decision problems of finite automata design and related arithmetics. *Transactions of the American Mathematical Society*, 98:21–51, 1961.

[ER66] C.C. Elgot and M.O. Rabin. Decidability and Undecidability of Extensions of Second (First) Order Theory of (Generalized) Successor. *Journal of Symbolic Logic*, 31:169–181, 1966.

[Grz05] Andrzej Grzegorczyk. Undecidability without arithmetization. *Studia Logica*, 79:163–230, 2005.

[Han65] W. Hanf. Model-theoretic methods in the study of elementary logic. In J.W. Addison, L. Henkin, and A. Tarski, editors, *The Theory of Models, Proceedings of the 1963 International Symposium at Berkeley*, pages 132–145. North Holland, Amsterdam, 1965.

[HP91] P. Hájek and P. Pudlák. *Metamathematics of First-Order Arithmetic.* Perspectives in Mathematical Logic. Springer, Berlin, 1991.

[Mac71] S. MacLane. *Categories for the Working Mathematician.* Number 5 in Graduate Texts in Mathematics. Springer, New York, 1971.

[MM94] F. Montagna and A. Mancini. A minimal predicative set theory. *The Notre Dame Journal of Formal Logic*, 35:186–203, 1994.

[MPS90] J. Mycielski, P. Pudlák, and A.S. Stern. *A lattice of chapters of mathematics (interpretations between theorems)*, volume 426 of *Memoirs of the American Mathematical Society.* AMS, Providence, Rhode Island, 1990.

[Pud83] P. Pudlák. Some prime elements in the lattice of interpretability types. *Transactions of the American Mathematical Society*, 280:255–275, 1983.

[Pud85] P. Pudlák. Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, 50:423–441, 1985.

[Smo85] C. Smoryński. Nonstandard models and related developments. In L.A. Harrington, M.D. Morley, A. Scedrov, and S.G. Simpson, editors, *Harvey Friedman's Research on the Foundations of Mathematics*, pages 179–229. North Holland, Amsterdam, 1985.

[Solle] R. M. Solovay. Interpretability in set theories. Unpublished letter to P. Hájek, 1976, Aug. 17. ⟨`http://www.cs.cas.cz/~hajek/RSolovayZFGB.pdf`⟩.

[ST50] W. Szmielew and A. Tarski. Mutual Interpretability of some essentially undecidable theories. In *Proceedings of the International Congress of Mathematicians*, volume 1, page 734. Cambridge, 1950.

[Ste89] A.S. Stern. Sequential theories and infinite distributivity in the lattice of chapters. *Journal of Philosophical Logic*, 54:190–206, 1989.

[Ten] R.L. Tenney. Decidable pairing functions. Unpublished.

[TMR53] A. Tarski, A. Mostowski, and R.M. Robinson. *Undecidable theories.* North–Holland, Amsterdam, 1953.

[Vis06] A. Visser. The predicative Frege hierarchy. Logic Group Preprint Series 246, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, ⟨`http://www.phil.uu.nl/preprints/lgps/`⟩, 2006.

[Vis07] A. Visser. Pairs, sets and sequences in first order theories. Logic Group Preprint Series 251, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, ⟨`http://www.phil.uu.nl/preprints/lgps/`⟩, 2007.