

On \mathbf{Q}

Albert Visser¹

Published online: 19 September 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract In this paper we study the theory \mathbf{Q} . We prove a basic result that says that, in a sense explained in the paper, \mathbf{Q} can be split into two parts. We prove some consequences of this result. (i) \mathbf{Q} is not a poly-pair theory. This means that, in a strong sense, pairing cannot be defined in \mathbf{Q} . (ii) \mathbf{Q} does not have the Pudlák Property. This means that there two interpretations of \mathbf{S}_2^1 in \mathbf{Q} which do not have a definably isomorphic cut. (iii) \mathbf{Q} is not sententially equivalent with \mathbf{PA}^- . This tells us that we cannot do much better than mutual faithful interpretability as a measure of sameness of \mathbf{Q} and \mathbf{PA}^- . We briefly consider the idea of characterizing \mathbf{Q} as the minimal-in-some-sense theory of some kind modulo some equivalence relation. We show that at least one possible road towards this aim is closed.

Keywords \mathbf{Q} · \mathbf{PA}^- · Arithmetic · Weak theories · Interpretability · comparison of theories

1 Prelude

Robinson's Arithmetic \mathbf{Q} is an old friend. The first time I met it was when I studied the First Incompleteness Theorem in Boolos and Jeffrey's wonderful book *Computability*

Communicated by A. Di Nola, D. Mundici, C. Toffalori, A. Ursini.

This paper is dedicated to the memory of Franco Montagna. Franco was a gentle colleague and a fine human being. His insights and ideas were always an inspiration.

✉ Albert Visser
a.visser@uu.nl

¹ Philosophy, Faculty of Humanities, Utrecht University, Janskerkhof 13, 3512 BL Utrecht, The Netherlands

and Logic.¹ The choice of \mathbf{Q} and its smaller brother \mathbf{R} to prove the *First Incompleteness Theorem* is beautiful, since these theories seem to be about as weak as one can get to prove this result.² On the other hand, it is a bit awkward to go on from there to prove the *Second Incompleteness Theorem*, where we need internal verification of some principles. *Computability and Logic* also contains nice exercises illustrating how easy it is to manufacture various counter-models to show non-provability in \mathbf{Q} .

Later I was pleased to discover that there are—in a sense—two proofs of the Second Incompleteness Theorem for \mathbf{Q} itself. There is a truly beautiful result by Bezboruah and Shepherdson (1976). They show that, under some very reasonable assumptions, the theory \mathbf{PA}^- does not prove the consistency of *any* theory. Hence, a fortiori, \mathbf{Q} does not prove its own consistency. Bezboruah and Shepherdson's proof does depend on rather specific assumptions about the coding. Also it does not generalize to stronger theories. What is more, it tells us nothing about the question whether \mathbf{Q} can prove its consistency on some definable cut. Later, Pavel Pudlák proved a

¹ *Caveat Emptor*. The version of *Computability and Logic* I studied was the second edition. In this book, \mathbf{Q} is introduced with the correct axioms but without the name 'Robinson's Arithmetic' and without reference to Tarski et al. (1953). The fourth edition (Boolos et al. 2002) is by Boolos, Burgess and Jeffrey. In this book, a different (but closely related) theory, also called *minimal arithmetic*, is presented as \mathbf{Q} . In an optional section, we learn that "the label \mathbf{Q} is often used to refer not to our minimal arithmetic but to another system, called *Robinson's Arithmetic* for which we use the label \mathbf{R} ". Then, the correct axioms for our \mathbf{Q} are specified. However, now the name \mathbf{R} is used for \mathbf{Q} and not for Tarski, Mostowski and Robinson's \mathbf{R} ...

² Saul Kripke noted that one can even prove the First Incompleteness Theorem for weaker theories than \mathbf{R} , to wit a theory he calls *school*. His proof uses Matiyasevich's Theorem to compensate the fact that for *school* one cannot prove Σ_1 -completeness, employing the fact that one still has completeness for purely existential formulas.

strong version of the Second Incompleteness Theorem (see Pudlák 1985; Hájek and Pudlák 1993) that I would formulate as follows. Let U be any consistent recursively enumerable theory and let N be an interpretation of \mathbf{Q} in U . Then, U does not prove its own consistency relativized to N . It seems to me that one should say that Bezboruah and Shepherdson on the one hand and Pudlák on the other proved quite different results which share some consequences. I always thought that the two proofs should provide a good case study for a philosophical enquiry into the problem of theorem individuation.

Still later I read Tarski et al. (1953) which proves the essential undecidability of \mathbf{Q} . And after that there was also Nelson (1986) where the possibilities for bootstrapping in \mathbf{Q} are used with impressive, almost magical, results.

The present paper is my tribute to \mathbf{Q} . As is fitting among friends, I will not only praise it, but also discuss, in a respectful way, some of its weaknesses.

2 Introduction

The theory \mathbf{Q} was introduced by Robinson (1950). He introduced it as a simplification of an earlier finitely axiomatized, essentially undecidable theory due to Mostowski and Tarski (1949). The system became widely known via the book Undecidable Theories by Tarski et al. (1953). It is given by the following axioms:

- Q1. $\vdash \mathbf{S}x = \mathbf{S}y \rightarrow x = y$
- Q2. $\vdash \mathbf{S}x \neq 0$
- Q3. $\vdash x = 0 \vee \exists y \ x = \mathbf{S}y$
- Q4. $\vdash x + 0 = x$
- Q5. $\vdash x + \mathbf{S}y = \mathbf{S}(x + y)$
- Q6. $\vdash x \cdot 0 = 0$
- Q7. $\vdash x \cdot \mathbf{S}y = x \cdot y + x$

Robinson shows that \mathbf{Q} is *essentially undecidable*, which tells us that any consistent theory that interprets \mathbf{Q} is undecidable. Since \mathbf{Q} is finitely axiomatized, it follows that any theory that weakly interprets \mathbf{Q} ,³ i.o.w. any theory that is consistent with some translation of \mathbf{Q} , is undecidable.

The theory \mathbf{Q} is, in many senses, a natural theory. What would be the quintessential weak arithmetic? Well, we want to have at least the basic properties of zero and the successor function: zero is not a successor and the successor relation is total, functional and injective. Then, we want the recursion equations for addition and multiplication. However, the resulting theory is a subtheory of the theory of the positive

reals including zero. Thus, it has a decidable extension. So, the theory cannot binumerate the recursive functions. We can repair that by adding the axiom that zero and successor are jointly surjective. This final axiom seems to have a certain harmony with the other successor axioms: we have axioms Q1 and Q2 to state that zero and successor are *jointly injective* and we have an axiom, to wit Q3, that articulates that they are *jointly surjective*. Thus, from the standpoint of motivation of the axioms, \mathbf{Q} seems to be a well-balanced theory.

Clearly, \mathbf{Q} is very weak. However, if we are just interested in the property of essential undecidability, we can go much weaker. There is the theory \mathbf{R} , also introduced in Tarski et al. (1953). Here are the axioms of \mathbf{R} , where underlining stands for the usual unary numeral function.

- R1. $\vdash \underline{m} + \underline{n} = \underline{m + n}$
- R2. $\vdash \underline{m} \cdot \underline{n} = \underline{m \cdot n}$
- R3. $\vdash \underline{m} \neq \underline{n}$, for $m \neq n$
- R4. $\vdash x \leq \underline{n} \rightarrow \bigvee_{i \leq n} x = \underline{i}$
- R5. $\vdash x \leq \underline{n} \vee \underline{n} \leq x$

This theory is not only essentially undecidable but also has the property that any theory that weakly interprets it is undecidable. See Vaught (1962). Since we build up the needed machinery anyway, we will provide a proof of this last result in Sect. 3.

We can even consider weaker theories than \mathbf{R} . See Vaught (1962) and Jones and Shepherdson (1983). See also Visser (2014) for more information.

The theory \mathbf{R} is not finitely axiomatizable. Moreover, every finitely axiomatized subtheory of it is finitely satisfiable and, hence, has a decidable extension. So perhaps \mathbf{Q} is the weakest *finitely axiomatized* theory that is essentially undecidable and, say, extends \mathbf{R} ? Alas, no such luck. We refer the reader to Vítězslav Švejdar's paper in Švejdar (2007) where finitely axiomatized systems are studied that are weaker than \mathbf{Q} . These systems do not extend \mathbf{R} but with a minor modification they do: we have to set the partial functions in Švejdar's paper when they are not defined to some default value.

The systems studied by Švejdar are mutually interpretable with \mathbf{Q} , so perhaps \mathbf{Q} is still minimal, among finitely axiomatized theories, with respect to interpretability? Not so, for any finitely axiomatized subtheory A of \mathbf{Q} that extends \mathbf{R} , we can find a finitely axiomatized subtheory B of A that extends \mathbf{R} and such that B does not interpret A . This can be shown by a minor adaptation of the methods of Friedman (2007). We show how to do this in Sect. 3. So, the prospect of characterizing \mathbf{Q} , or some closely related theory, as the weakest (in a suitable sense) finitely axiomatized theory with such and such a property, seems pretty dim.

A quite different and beautiful feature of \mathbf{Q} is that it interprets fairly strong theories like $I\Delta_0 + \Omega_1$ on a definable cut. It follows from this that we have the second incompleteness

³ The notion of *weak interpretability* is introduced by Tarski, Mostowski and Robinson. Giorgi Japaridze uses *tolerates* for *weakly interprets*, which seems to me a more descriptive term.

theorem for all extensions of \mathbf{Q} . This feature also holds for the still weaker theories studied by Švejdar that interpret \mathbf{Q} on a definable cut. Regrettably, this does not seem to help us with the characterization problem.

From one perspective, \mathbf{Q} seems rather natural, from another it does not. It lacks many desirable properties. As has been shown in Visser (2008) and Jeřábek (2012), \mathbf{Q} is not a pair theory. In Sect. 5 of this paper we will show that it is not even a poly-pair theory. We will also show, in Sect. 6, that \mathbf{Q} does not have the Pudlák property. The negation of the Pudlák property for \mathbf{Q} tells us that there are two interpretations of \mathbf{S}_2^1 in \mathbf{Q} that do not verifiably have definably isomorphic cuts.

It is interesting to compare \mathbf{Q} with its bigger brother \mathbf{PA}^- . The theory \mathbf{PA}^- is the theory of commutative, discretely ordered semi-rings with a minimal element plus the subtraction axiom (\mathbf{PA}^- 14 below). It is employed as the basic arithmetic, e.g. in the textbook (Kaye 1991). The theory is given by the following axioms:

- \mathbf{PA}^- 1. $\vdash x + 0 = x$
- \mathbf{PA}^- 2. $\vdash x + y = y + x$
- \mathbf{PA}^- 3. $\vdash (x + y) + z = x + (y + z)$
- \mathbf{PA}^- 4. $\vdash x \cdot 1 = x$
- \mathbf{PA}^- 5. $\vdash x \cdot y = y \cdot x$
- \mathbf{PA}^- 6. $\vdash (x \cdot y) \cdot z = x \cdot (y \cdot z)$
- \mathbf{PA}^- 7. $\vdash x \cdot (y + z) = x \cdot y + x \cdot z$
- \mathbf{PA}^- 8. $\vdash x \leq y \vee y \leq x$
- \mathbf{PA}^- 9. $\vdash (x \leq y \wedge y \leq z) \rightarrow x \leq z$
- \mathbf{PA}^- 10. $\vdash x + 1 \not\leq x$
- \mathbf{PA}^- 11. $\vdash x \leq y \rightarrow (x = y \vee x + 1 \leq y)$
- \mathbf{PA}^- 12. $\vdash x \leq y \rightarrow x + z \leq y + z$
- \mathbf{PA}^- 13. $\vdash x \leq y \rightarrow x \cdot z \leq y \cdot z$
- \mathbf{PA}^- 14. $\vdash x \leq y \rightarrow \exists z \ x + z = y$

Emil Jeřábek's version in Jeřábek (2012) does not have the subtraction axiom \mathbf{PA}^- 14. Thus Jeřábek's version is a universal theory. As noted by Jeřábek his version interprets the stronger version with subtraction axiom on a definable cut. The weak version does not extend \mathbf{Q} but the strong version does.

The theory \mathbf{PA}^- has been shown to be sequential by Emil Jeřábek in his paper (Jeřábek 2012) (even in the weaker form without the subtraction axiom).

Bezboruah and Shepherdson (1976) show that, under some very reasonable assumptions, \mathbf{PA}^- does not prove the consistency of *any* theory.

Victor Pambuccian studies number theoretical theorems over \mathbf{PA}^- . See his papers (Pambuccian 2008, 2014, 2015).

The theory \mathbf{Q} interprets much stronger theories than \mathbf{PA}^- , like $I\Delta_0 + \Omega_1$ (see, e.g. Hájek and Pudlák 1993). Hence, a fortiori, \mathbf{Q} is mutually interpretable with \mathbf{PA}^- (both the strong and the weak version). Using ideas of Per Lindström, one may

show that \mathbf{Q} is even mutually *faithfully* interpretable with \mathbf{PA}^- . One can also demonstrate that the Lindenbaum algebras of \mathbf{Q} and \mathbf{PA}^- are recursively isomorphic. This means that there is a recursive function of sentences that induces an isomorphism of Lindenbaum algebras. The result is a special case of the theorem of Marian Pour-El and Saul Kripke that the Lindenbaum algebras of all recursively enumerable theories that interpret \mathbf{Q} are recursively isomorphic. See Pour-El and Kripke (1967).

Are these the best samenesses that we can get between these theories? In Sect. 7, we will show that the two theories are not *sententially congruent*. So, at least in terms of traditional notions of sameness, we cannot do better than recursive isomorphism of Lindenbaum algebras on the one hand, and mutual faithful interpretability on the other.

The main technical tool of the paper is a theorem that tells us that, in a sense, \mathbf{Q} can be split into two disjoint parts. This result is proved in Sect. 4. The proof is an adaptation of an earlier result in Visser (2014). One might say that the progress of the present paper is to provide a better understanding of what the result of Visser (2014) really means.

We end the paper with some concluding remarks in Sect. 8.

2.1 How to read the paper

In the appendices, I present basic materials needed for understanding the paper. In the main text there are references to the appendices when needed. Section 3 can be read independently of the other sections. Section 4 is the basic preliminary for Sects. 5–7. The Sects. 5–7 are pairwise independent of each other.

3 Between R and Q

In this section, we endeavour to make the idea of characterizing \mathbf{Q} using an appropriate minimality claim less plausible. Perhaps it is better to say: if there is a characterization of \mathbf{Q} as the minimal theory such that ..., then it cannot take such and such a form. Specifically, we show that, for any finitely axiomatised consistent theory A such that $\mathbf{R} \subseteq A$, there is a finitely axiomatised B such that $\mathbf{R} \subseteq B \subseteq A$ and $B \not\vdash A$.⁴ The result is just a rather direct application of ideas from Friedman (2007). So, we do not claim great originality here.

The following nice version of the theory of a number was developed by Johannes Marti, Nal Kalchbrenner, Paula Henk and Peter Fritz in *Interpretability Project Report* of 2011, the report of a project they did under my guidance in the Master

⁴ The notation \subseteq holds between theories of the same signature. It is the extensional subset relation between the sets of theorems of the theories at hand. The notation \triangleright is explained in the "Global and local interpretability" section of Appendix 1.

of Logic in Amsterdam.^{5,6} We call it the theory of a number since, in our intended applications, the fact that it is satisfied by the structure associated with a finite, non-zero ordinal is central.

- TN1. $\vdash x \neq 0$
- TN2. $\vdash (x < y \wedge y < z) \rightarrow x < z$
- TN3. $\vdash x < y \vee x = y \vee y < x$
- TN4. $\vdash x = 0 \vee \exists y \quad x = \mathbf{S}y$
- TN5. $\vdash \mathbf{S}x \neq x$
- TN6. $\vdash x < y \rightarrow (x < \mathbf{S}x \wedge y \neq \mathbf{S}x)$
- TN7. $\vdash x + 0 = 0$
- TN8. $\vdash x + \mathbf{S}y = \mathbf{S}(x + y)$
- TN9. $\vdash x \cdot 0 = 0$
- TN10. $\vdash x \cdot \mathbf{S}y = x \cdot y + x$

Since TN6 implies $x \neq x$, a model of TN is a linear ordering that either represents a finite ordinal or starts with a copy of ω .

We call a Δ_0 -formula *pure* if (i) all bounding terms are variables and (ii) all occurrences of terms are in subformulas of the form $\mathbf{S}x = y$, $x + y = z$ and $x \cdot y = z$. We call a Σ_1 -sentence *pure* if it is of the form $\exists x S_0x$, where S_0 is a pure Δ_0 -sentence.

We can transform an arbitrary Σ_1 -sentence S into a pure Σ_1 -sentence S° , for example, in the following way. We start with S . We treat bounded quantifiers for the moment as if they were given with the language and not defined. First we replace all implications ($A \rightarrow B$) by $(\neg A \vee B)$ and all bi-implications by $((\neg A \vee B) \wedge (\neg B \vee A))$. Next we push all negations inside in the usual manner. We replace:

- $\forall x < t$ by $\exists z \quad (t = z \wedge \forall x < z \dots)$,
- $\exists x < t$ by $\exists z \quad (t = z \wedge \exists x < z \dots)$,
- $\neg t_0 = t_1$ by $\exists z \exists w \quad (t_0 = z \wedge t_1 = w \wedge \neg z = w)$,
- $\neg t_0 < t_1$ by $\exists z \exists w \quad (t_0 = z \wedge t_1 = w \wedge \neg z < w)$.

In this way all term occurrences are on positive places. At this point we apply the usual term-unwinding algorithm to our formula using a small scope interpretation. We note that this will translate an atomic formula to a block of existential quantifiers followed by a boolean combination of atomic formulas. Finally, we bring all unbounded existential quantifiers to the front in the usual manner replacing, e.g. $\forall x < y \exists z \dots$ by $\exists w \forall x < y \exists z < w \dots$. The resulting

⁵ Unfortunately, the report is not available. We reproduce all necessary materials here.

⁶ I simplified the axioms of Marti, Kalchbrenner, Henk and Fritz a bit and also implemented three nice simplifications suggested by the referee. It seems to me that, for our intended applications, we could even work with the weaker system that omits axiom TN4. The remaining system would also be satisfied by an arbitrary ordinal.

formula is S° , which is clearly pure and equivalent to the original formula (say, over PA^- plus Σ_1 -collection).⁷

We note that the transformation $S \mapsto S^\circ$ that we described is clearly elementary. Hence it exists in EA. Inspecting the transformation, we see that S° implies S in predicate logic.

Let $S := \exists y S_0y$, where S_0 is a pure Δ_0 -formula. We define:

$$[S] := \text{TN} + \exists x \exists y < x S_0y.$$

Using the machinery of theories of a number, we can reprove Cobham's result that any recursively enumerable theory that weakly interprets \mathbf{R} is undecidable. Suppose U is recursively enumerable and $U + \mathbf{R}^\tau$ is consistent. Consider a pure Σ_1 -sentence S . Let S^* be the sentence that says:

$$\exists x (\mathbf{S}x \models [S] \wedge \forall y < x \mathbf{S}y \not\models [S]).$$

Since $[S]$ is finitely axiomatised, we can write out S^* in the obvious way. Consider the set \mathcal{S} of all S such that $U + (\mathbf{R} + S^*)^\tau$ is consistent. Clearly, \mathcal{S} contains all true (pure) Σ_1 -sentences. If \mathcal{S} did not contain false (pure) Σ_1 -sentences, then this would make Σ_1 -truth decidable. Hence, there is a false (pure) Σ_1 -sentence S_1 such that $U + (\mathbf{R} + S_1^*)^\tau$ is consistent. We can use S_1 to build a translation τ_0 such that $U + [S_1]^{\tau_0}$ is consistent. Since $[S_1]$ is a finitely axiomatised extension of \mathbf{R} , it follows that U is undecidable.

To prove our main result we need the following result that was first verified in detail in the *Interpretability Project Report* by Marti, Kalchbrenner, Henk and Fritz. The basic idea behind the result is present in Friedman's (2007).

We remind the reader of witness comparison notation. Suppose A is of the form $\exists x A_0(x)$ and B is of the form $\exists y B(y)$. We define:

- $A < B := \exists x (A(x) \wedge \forall y \leq x \neg B(y))$.
- $A \leq B := \exists x (A(x) \wedge \forall y < x \neg B(y))$.
- If C is $A < B$, then C^\perp is $B \leq A$.
- If D is $A \leq B$, then D^\perp is $B < A$.

Theorem 1 *Let S and S' be pure Σ_1 -sentences. We have:*

- (a) *Suppose S is true. Then, if we allow piecewise interpretations, we have $\top \triangleright [S]$.*

If we do not allow piecewise interpretations, we still have $(\exists x \exists y x \neq y) \triangleright [S]$.

- (b) *If $S \leq S'$, then $[S'] \vdash S$.*

⁷ The referee remarks that the use of Σ_1 -collection can be eliminated from the argument. To do this one needs a careful expansion and reworking of the current argument.

Proof Ad (a): We note that if S is true, then $[S]$ has a finite model. Any theory with a finite model is interpretable with a piecewise interpretation in predicate logic. If we do not allow piecewise interpretations, we can obtain the same effect using a multidimensional interpretation, assuming that we have at least two distinct elements.

Ad (b). Consider any model \mathcal{M} of $[S']$. Without loss of generality we can identify the initial elements of \mathcal{M} with $0, 1, 2, \dots$. It is easily shown that for any pure Δ_0 -formula Ax we have $A\mathbf{n}$ is true iff $\mathcal{M} \models A\mathbf{n}$, provided the \mathbf{n} are natural numbers in \mathcal{M} and are not the top elements.⁸ Suppose m is the smallest witness of S . Then we have $\mathcal{M} \models \neg S'_0\mathbf{k}$, for all $\mathbf{k} < m$. Since S' is witnessed by a non-top element, m is non-top and we have $\mathcal{M} \models S_{0m}$, and hence $\mathcal{M} \models S$. \square

We now have the materials to prove the main theorem of the present section.

Theorem 2 *Suppose $R \subseteq A$, where A is finitely axiomatized and consistent. Then, there is a finitely axiomatized B such that $R \subseteq B \subseteq A$ and $B \not\vdash A$.*

Proof Suppose $R \subseteq A$, where A is finitely axiomatized and consistent. By the Gödel Fixed Point Lemma, we define R with:

$$EA \vdash R \leftrightarrow ([R^\circ] \triangleright A) \leq (A \triangleright [R^\circ]).$$

Suppose $[R^\circ] \triangleright A$. It follows that either R or R^\perp .

In the first case, we find R° . By Theorem 1(a), we find that $\top \triangleright [R^\circ]$ and, hence, that $\top \triangleright A$. *Quod non*, since A extends R and since any theory interpretable in predicate logic has finite models.

In the second case, it follows that $(R^\perp)^\circ$ and $\neg R^\circ$. Hence, $(R^\perp)^\circ \leq R^\circ$ and, hence, by Theorem 1(b), $[R^\circ] \vdash (R^\perp)^\circ$. Since R° implies R in predicate logic and similarly for $(R^\perp)^\circ$ and R^\perp , we find that $[R^\circ]$ proves both R and R^\perp . So, $[R^\circ] \vdash \perp$. From R^\perp , we also have $A \triangleright [R^\circ]$, thus it follows that $A \triangleright \perp$. *Quod non*, since A is consistent.

We may conclude that $[R^\circ] \not\vdash A$ and hence that $(\bigwedge [R^\circ] \vee A) \not\vdash A$. Since R implies $[R^\circ] \triangleright A$, it also follows that R is false. Thus, $[R^\circ]$ extends R . Let $B := (\bigwedge [R^\circ] \vee A)$. We find: $R \subseteq B \subseteq A$ and $B \not\vdash A$. \square

4 Decomposition of Q

The theory Q has many unexpected extensions. We develop one here that is especially useful for proving negative results about Q . We will use it in the subsequent sections. The construction is an adaptation of a result of Visser (2014). One

⁸ Note that the \mathbf{n} in the context of $\mathcal{M} \models A\mathbf{n}$ are used as sequences of domain constants and not as sequences of numerals.

could say that only in the present paper the full meaning of the construction of Visser (2014) is unfolded.

The theory $Q^\#$ in the language of Q is axiomatized by the following principles.

- Q
- $\forall x \forall y ((Sx \neq x \wedge Sy \neq y) \rightarrow S(x + y) \neq x + y)$
- $\forall x \forall y ((Sx \neq x \wedge Sy \neq y) \rightarrow S(x \cdot y) \neq x \cdot y)$
- $\exists a \forall x (x + a = a \wedge a + x = a)$
- It is easy to see that this a is unique. We call it ∞ .
- $\exists x (Sx = x \wedge x \neq \infty)$
- $\forall x \forall y ((x \neq Sx \wedge y = Sy) \rightarrow (x + y = y \wedge y + x = y))$
- $\forall x \forall y ((x = Sx \wedge y = Sy) \rightarrow (x + y = y \vee x + y = \infty))$
- $\forall x \forall y (y = Sy \rightarrow x \cdot y = \infty)$
- $\forall x \forall y ((x = Sx \wedge y \neq Sy) \rightarrow x \cdot SSy = x + x)$

Theorem 3 below will have the immediate consequence that $Q^\#$ is consistent.

Let Q° be Q plus the axiom $\forall x Sx \neq x$. Let CQC^2 be predicate logic with identity and one binary predicate symbol R . Let $\mathbb{1}$ be the theory in the language of identity that states that there is at most one object. The operation \boxplus is defined and discussed in the ‘‘Sums’’ section of Appendix 1. The notion of *synonymy* is defined in the ‘‘Provable equivalence of interpretations’’ section of Appendix 1. We have:

Theorem 3 *The theory $Q^\#$ is synonymous to the theory $Y := Q^\circ \boxplus CQC^2 \boxplus \mathbb{1}$.*

Proof We define $K : Y \rightarrow Q^\#$ as follows. We write Z for the unary relational representation of zero, S for the binary relational representation of successor, A for the ternary relational representation of addition, and M for the ternary relational representation of multiplication.

- $\delta_K(x) : \leftrightarrow x = x$
- $x =_K y : \leftrightarrow x = y$
- $\Delta_{0K}(x) : \leftrightarrow Sx \neq x$
- $\Delta_{1K}(x) : \leftrightarrow Sx = x \wedge x \neq \infty$
- $\Delta_{2K}(x) : \leftrightarrow x = \infty$
- $Z_K x : \leftrightarrow Zx$
- $S_K xy : \leftrightarrow \Delta_{0K}(x) \wedge Sx = y$
- $A_K xyz : \leftrightarrow \Delta_{0K}(x) \wedge \Delta_{0K}(y) \wedge x + y = z$
- $M_K xyz : \leftrightarrow \Delta_{0K}(x) \wedge \Delta_{0K}(y) \wedge x \cdot y = z$
- $R_K xy : \leftrightarrow \Delta_{1K}(x) \wedge \Delta_{1K}(y) \wedge x + y = \infty$

It is easy to see that the specified translation does indeed deliver the desired interpretation. We define $M : Q^\# \rightarrow Y$ as follows. To make the interpretation readable we use functional notation on the side of the interpreting theory. The reader should keep in mind that, e.g. S is only defined on Δ_0 . We write ∞ for the unique inhabitant of Δ_2 . Below the itemized definition, we repeat the definitions of addition and multiplication in more readable tabular form.

- $\delta_M(x) :\leftrightarrow x = x$
- $x =_M y :\leftrightarrow x = y$
- $\mathbf{Z}_M x :\leftrightarrow x = 0$
- $\mathbf{S}_M xy :\leftrightarrow \mathbf{S}x = y \vee (\neg\Delta_0(x) \wedge x = y)$
- $\mathbf{A}_M xyz :\leftrightarrow x + y = z \vee$
 $(\Delta_0(x) \wedge \Delta_1(y) \wedge z = y) \vee$
 $(\Delta_1(x) \wedge \Delta_0(y) \wedge z = x) \vee$
 $(\Delta_1(x) \wedge \Delta_1(y) \wedge Rxy \wedge z = \infty)$
 $(\Delta_1(x) \wedge \Delta_1(y) \wedge \neg Rxy \wedge z = y) \vee$
 $((x = \infty \vee y = \infty) \wedge z = \infty)$
- $\mathbf{M}_M xyz :\leftrightarrow (y = 0 \wedge z = 0) \vee (y = 1 \wedge \mathbf{A}_M(0, x, z)) \vee$
 $\exists u(y = \mathbf{S}Su \wedge ((\Delta_0(x) \wedge x \cdot y = z) \vee$
 $(\Delta_1(x) \wedge ((Rxx \wedge z = \infty) \vee (\neg Rxx \wedge z = x)))) \vee$
 $(x = \infty \wedge z = \infty))) \vee ((\Delta_1(y) \vee y = \infty) \wedge z = \infty)$

Here is the diagrammatic version of the definitions of addition and multiplication. In the diagrams, n ranges over Δ_0 and x ranges over Δ_1 .

$+_M$	n	y	∞
m	$m+n$	y	∞
x	x	∞ if $R(x, y)$ y otherwise	∞
∞	∞	∞	∞

\cdot_M	0	1	$n \neq 0, 1$	y	∞
m	0	$0+m$	$m \cdot n$	∞	∞
x	0	x	∞ if $R(x, x)$ x otherwise	∞	∞
∞	0	∞	∞	∞	∞

The verification that the translation we specified does indeed carry an interpretation of $\mathbf{Q}^\#$ is immediate. We treat two sample cases of the verification of $M \circ K =_0 \text{ID}_Y$. We have:

$$\begin{aligned} Y \vdash (\mathbf{A}xyz)^{KM} &\leftrightarrow (\mathbf{S}x \neq x \wedge \mathbf{S}y \neq y \wedge x + y = z)^M \\ &\leftrightarrow \Delta_0(x) \wedge \Delta_0(y) \wedge (\mathbf{A}xyz \vee (\dots)) \\ &\leftrightarrow \mathbf{A}xyz \end{aligned}$$

The last step uses the fact that the (\dots) implies that $\neg(\Delta_0(x) \wedge \Delta_0(y))$.

$$\begin{aligned} Y \vdash (\mathbf{R}xy)^{KM} &\leftrightarrow (\mathbf{S}x = x \wedge \mathbf{S}y = y \wedge x \neq \infty \wedge \\ &\quad y \neq \infty \wedge x + y = \infty)^M \\ &\leftrightarrow \Delta_1(x) \wedge \Delta_1(y) \wedge Rxy \\ &\leftrightarrow Rxy \end{aligned}$$

We treat one sample case to illustrate that $K \circ M =_0 \text{ID}_{\mathbf{Q}^\#}$.

$$\begin{aligned} \mathbf{Q}^\# \vdash (x + y = z)^{MK} &\leftrightarrow (x + y = z \vee \\ &\quad (\Delta_0(x) \wedge \Delta_1(y) \wedge z = y) \vee \\ &\quad (\Delta_1(x) \wedge \Delta_0(y) \wedge z = x) \vee \\ &\quad (\Delta_1(x) \wedge \Delta_1(y) \wedge Rxy \wedge z = \infty) \\ &\quad (\Delta_1(x) \wedge \Delta_1(y) \wedge \neg Rxy \wedge z = y) \vee \\ &\quad ((x = \infty \vee y = \infty) \wedge z = \infty))^K \\ &\leftrightarrow (\mathbf{S}x \neq x \wedge \mathbf{S}y \neq y \wedge x + y = z) \vee \\ &\quad (\mathbf{S}x \neq x \wedge \mathbf{S}y = y \wedge y \neq \infty \wedge z = y) \vee \\ &\quad (\mathbf{S}x = x \wedge x \neq \infty \wedge \mathbf{S}y \neq y \wedge z = x) \vee \\ &\quad (\mathbf{S}x = x \wedge x \neq \infty \wedge \mathbf{S}y = y \wedge y \neq \infty \wedge \\ &\quad x + y = \infty \wedge z = \infty) \\ &\quad (\mathbf{S}x = x \wedge x \neq \infty \wedge \mathbf{S}y = y \wedge y \neq \infty \wedge \\ &\quad x + y \neq \infty \wedge z = y) \vee \\ &\quad ((x = \infty \vee y = \infty) \wedge z = \infty) \\ &\leftrightarrow x + y = z \end{aligned}$$

Of course, the last step is by verifying that if, e.g. $\mathbf{S}x = x \wedge x \neq \infty \wedge \mathbf{S}y = y \wedge y \neq \infty$, then $x + y = \infty \wedge z = \infty$ iff $x + y = z$, and similarly for the other cases. \square

Theorem 4 $\mathbf{Q} \triangleright \mathbf{Q}^\#$.

Proof The theorem is a direct consequence of the fact that $\mathbf{Q} \triangleright \mathbf{Q}^\circ$ via a definable cut, $\mathbf{Q} \triangleright \text{CQC}^2$ and $\mathbf{Q} \triangleright \mathbb{1}$ in combination with Theorem 3, noting that \boxplus is (an implementation of) the supremum in the degrees of interpretability. \square

5 Q is not a poly-pair theory

In this section we show that \mathbf{Q} is not a poly-pair theory. We explain the notion of poly-pair theory in Appendix 2. In this same appendix we provide various basic facts about poly-pair theories. These basics are mainly derived from our paper (Visser 2013).

Theorem 5 \mathbf{Q} is not a poly-pair theory.

Proof Suppose \mathbf{Q} were a poly-pair theory. By our result of Sect. 4, the theory \mathbf{Y} is bi-interpretable with $\mathbf{Q}^\#$. Since, by the results of Appendix 2, the property of being a poly-pair theory is upwards preserved under theory extension and is preserved under bi-interpretations, is sufficient to show that \mathbf{Y} is not a poly-pair theory.

Suppose \mathbf{Y} is a poly-pair theory. Consider any model \mathcal{M} of \mathbf{Y} in which the relation R is empty and in which the domain of the second component is infinite. By the results in Appendix 2 the interpretation that witnesses the fact that \mathbf{Y} is poly-pair can be taken to be parameter-free, but we do not need to use

the result. Suppose the parameters of our interpretation are \mathbf{p} and let the dimension be m .

Consider two m -sequences \mathbf{a} and \mathbf{b} in the second component. We assume that the elements of \mathbf{a} , \mathbf{b} , \mathbf{p} are pairwise distinct. Let \mathbf{c} be a pair (according to the interpretation) containing \mathbf{a} and \mathbf{b} . Some element d of \mathbf{a} , \mathbf{b} does not occur in \mathbf{c} . Let e in the second component be disjoint from \mathbf{a} , \mathbf{b} , \mathbf{c} , \mathbf{p} . Let σ be the operation of interchanging d and e . Clearly, σ is an automorphism of our model that leaves \mathbf{p} fixed. So \mathbf{c} is also a pair of $\sigma\mathbf{a}$ and $\sigma\mathbf{b}$. Since either $\sigma\mathbf{a} \neq \mathbf{a}$ or $\sigma\mathbf{b} \neq \mathbf{b}$, this contradicts the defining property of pairing. \square

6 The Pudlák property

The Pudlák property of a theory U in its classical formulations says: (i) there is an interpretation $N^* : \mathbf{S}_2^1 \rightarrow U$ and (ii) whenever $N : \mathbf{S}_2^1 \rightarrow U$ and $N' : \mathbf{S}_2^1 \rightarrow U$, then there is a U -definable, U -verifiable isomorphism F between certain U -definable, U -verifiable cuts I of N and J of N' . We take our cuts to be downwards closed w.r.t. $<$ and closed under \mathbf{S} , $+$, \times and ω_1 .

To keep our treatment simple we only consider the case of parameter-free interpretations. The case with parameters is briefly discussed in Remark 1.

For our purposes, it is nicer to view the Pudlák property in the light of the category INT_1 of bi-interpretability. See Appendix 1 for an introduction to bi-interpretability and INT_1 .

We define a functor \mathbf{A} from INT_1 to the category of preorders. Here we allow the empty preorder. Consider any theory U . We send U to the structure $\mathbf{A}(U)$. The elements of $\mathbf{A}(U)$ are interpretations $N : \mathbf{S}_2^1 \rightarrow U$ modulo i -isomorphism. The structure $\mathbf{A}(U)$ has a binary preordering \leq defined by $N \leq N'$ iff there is a U -definable, U -verifiable initial embedding F from N to N' . We note that the existence of such an embedding is independent of the choice of the syntactical representatives of N and N' . If $K : U \rightarrow V$, then $\mathbf{A}(K) : \mathbf{A}(U) \rightarrow \mathbf{A}(V)$ is defined by $\mathbf{A}(K)(N) := K \circ N$. We note $\mathbf{A}(K)$ does indeed preserve \leq .

We remind the reader that a preorder is *downward directed* if for every x and y , there is a z with $z \leq x$ and $z \leq y$.

The Pudlák property for U is equivalent to: $\mathbf{A}(U)$ is non-empty and downward directed.

We remind the reader of Pavel Pudlák’s well-known result from Pudlák (1985).

Theorem 6 (Pudlák) *Sequential theories have the Pudlák property.*

We show that the converse of Pudlák’s result does not hold in Appendix 3. Specifically, we show that, if U has the Pudlák Property, then so does $U \boxplus \text{EQ}$. Here EQ is the pure theory of equality in the minimal signature. It follows that,

e.g. $\mathbf{S}_2^1 \boxplus \text{EQ}$ has the Pudlák Property. However, the methods of Sect. 7 show that $\mathbf{S}_2^1 \boxplus \text{EQ}$ is not sequential.

Remark 1 What is the Pudlák Property in the case with parameters? It seems that there are lots of options. Instead of systematically looking at various versions, I will just give the one that I think is most attractive. For the basic definitions and notations concerning parameters the reader is referred to the “Adding parameters” section of Appendix 1.

First we generalize \leq to the case with parameters. We define:

- $N \leq N'$ iff, for some F , we have:

$$U \vdash \forall \mathbf{q} (\alpha_{N'}(\mathbf{q}) \rightarrow \exists \mathbf{p} (\alpha_N(\mathbf{p}) \wedge F(\mathbf{p}, \mathbf{q}) : N^{\mathbf{p}} \leq N'^{\mathbf{q}})).$$

Here $(F(\mathbf{p}, \mathbf{q}) : N^{\mathbf{p}} \leq N'^{\mathbf{q}})$ means that F is a formula representing an initial embedding of $N^{\mathbf{p}}$ in $N'^{\mathbf{q}}$, where domain and range of F are cuts. We note that we could allow F to have some extra parameters of its own, so that the formula $(F(\mathbf{p}, \mathbf{q}) : N^{\mathbf{p}} \leq N'^{\mathbf{q}})$ would become $\exists \mathbf{r} (F(\mathbf{p}, \mathbf{q}, \mathbf{r}) : N^{\mathbf{p}} \leq N'^{\mathbf{q}})$. However, we will refrain from doing so.

The Pudlák Property now takes a simple form: (i) There is an $N^* : \mathbf{S}_2^1 \rightarrow U$ and (ii) for every $N : \mathbf{S}_2^1 \rightarrow U$, there is a parameter-free $N_0 : \mathbf{S}_2^1 \rightarrow U$, such that $N_0 \leq N$. This holds even in the case that the direct interpretation that witnesses the sequentiality of U itself contains parameters. We can rephrase this version of the Pudlák property as follows: $\mathbf{A}(U)$ is not empty, and the parameter-free interpretations are coinital in $\mathbf{A}(U)$.

We note that the two interpretations N and N' in the formulation of the parameter-free case can be subsumed under a single interpretation with parameters $N(x = 0)N'$. So, the Pudlák Property with parameters includes the one without parameters.

Let us call my version of the Pudlák Property with parameters: *the strong Pudlák Property*. Sequential theories have the strong Pudlák Property. This result holds even in the case that U is sequential via a direct interpretation with parameters. The result also holds in the poly-sequential case. We refer the reader to Visser (2013), for details, specifically to the second proof of Theorem 5.2 of that paper.

I have not worked out the full development of the case with parameters. However note that we show that \mathbf{Q} fails to have the weaker property, so a fortiori it fails to have the stronger property.

We collect some basic insights. Since the homomorphic image of a downward directed preorder is downward directed, we have:

Theorem 7 *Suppose $K : U \rightarrow V$ and U has the Pudlák Property and $\mathbf{A}(K)$ is surjective. Then, V has the Pudlák property.*

We show that \mathbf{A} applied to an instance of the extension relation is surjective.

Theorem 8 *Suppose V is an extension of U in the same language. Let emb_{UV} be the identical embedding. Suppose further that $\mathbf{A}(U)$ is non-empty. Then, $\mathbf{A}(\text{emb}_{UV})$ is surjective.*

Proof Suppose $N_0 \in \mathbf{A}(U)$ and $N \in \mathbf{A}(V)$. Let $N_1 := N \langle (\bigwedge \mathbf{S}_2^1)^N \rangle N_0$.⁹ Then, clearly $N_1 \in \mathbf{A}(U)$. Moreover, $\text{emb}_{UV}(N_1) = N$. \square

We remind the reader of the following. Consider a category \mathcal{C} . Suppose $f : x \rightarrow y$ and $g : y \rightarrow x$ and $g \circ f = \text{id}_x$. In this case, we call f a *section* or *split monomorphism*. We call g a *retraction* or *split epimorphism*. The object x is in this situation a *retract* of y .

We show that \mathbf{A} applied to a retraction is surjective.

Theorem 9 *Suppose $K : U \rightarrow V$ is a retraction, then $\mathbf{A}(K)$ is a retraction and hence surjective.*

Proof Any functor preserves retractions. So $\mathbf{A}(K)$ is a retraction in the category of preorders. It follows that $\mathbf{A}(K)$ is surjective. \square

Question 1 We note that both retractions and theory extensions are epimorphisms in INT_1 . Does \mathbf{A} preserve epimorphisms? (Clearly an epimorphism in the category of preorders is surjective.)

In Visser (2006), we proved that, in INT_0 , epimorphisms can always be split in first a theory extension and then an isomorphism. So, a fortiori, in INT_0 , epimorphisms are preserved by the INT_0 -analogue of \mathbf{A} .

Theorem 10 *\mathbf{Q} does not have the Pudlák Property.*

Proof By Theorem 3, we have extensions $\mathbf{Q}^\#$ and \mathbf{Q}° of \mathbf{Q} such that $\mathbf{Q}^\#$ is synonymous to $\mathbf{Y} := \mathbf{Q}^\circ \boxplus \text{CQC}^2 \boxplus \mathbb{1}$. We extend CQC^2 to AS with R in the role of \in .¹⁰ Let $A := \mathbf{Q}^\circ \boxplus \text{AS} \boxplus \mathbb{1}$.

Suppose \mathbf{Q} has the Pudlák Property. By Theorems 8 and 9, the property is preserved from \mathbf{Q} to $\mathbf{Q}^\#$, from $\mathbf{Q}^\#$ to \mathbf{Y} , and from \mathbf{Y} to A .

Consider interpretations $N : \mathbf{S}_2^1 \rightarrow \mathbf{Q}^\circ$ and $M : \mathbf{S}_2^1 \rightarrow \text{AS}$. Let $N^* := \text{in}_0 \circ N : \mathbf{S}_2^1 \rightarrow A$ and let $M^* := \text{in}_1 \circ M : \mathbf{S}_2^1 \rightarrow A$. Suppose there is an embedding F in A of a cut of N^* into a cut of M^* . By Theorem 15, we have:

$$A, \bigwedge_i \Delta_0(x_i), \bigwedge_\ell \Delta_1(y_\ell) \vdash F(\mathbf{x}, \mathbf{y}) \leftrightarrow \bigvee_j (D_j(\mathbf{x}) \wedge E_j(\mathbf{y})).$$

⁹ The notation $K \langle A \rangle M$ is explained in the ‘‘Translations’’ and ‘‘Relative interpretations’’ sections of Appendix 1.

¹⁰ For the definition of AS , see Appendix 2.

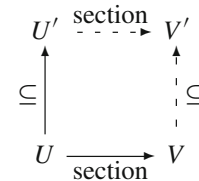
Here the D_j are formulas in the range of in_0 and the E_j are formulas in the range of in_1 . Suppose \mathbf{x} is in D_j . In that case all \mathbf{y} in E_j are in the F -image of \mathbf{x} . Hence, E_j is closed under $=_{M^*}$. We may conclude that the range of F is standardly finite modulo $=_{M^*}$. Quod non. \square

7 PA^- and \mathbf{Q}

In this section we show that PA^- is not sententially congruent with \mathbf{Q} .¹¹ In a sense, we could have written this section without even mentioning PA^- , since the result that PA^- and \mathbf{Q} are not sententially congruent follows from a much stronger result proven here. However, since the theories PA^- and \mathbf{Q} seem to be so close together, I feel that the specific result concerning PA^- and \mathbf{Q} speaks more to the imagination than the stronger result from which it follows.

To prove our main result, we need a few lemmas. We will be interested in retractions in the category INT_3 (see the ‘‘Five categories’’ section of Appendix 1). This takes the following form: we have interpretations $K : U \rightarrow V$ and $M : V \rightarrow U$ such that, for all U -sentences A , $U \vdash A \leftrightarrow A^{KM}$. In this case K is a section or split monomorphism.

A basic insight is that the section relation has the forward or zig property w.r.t. theory-extension in INT_3 . This is illustrated by the following diagram.



Theorem 11 *The section relation in INT_3 has the forward or zig property with respect to theory extension.*

Proof Suppose $K : U \rightarrow V$ is a section. We suppose that M is an inverse of K , so $M : V \rightarrow U$ and $M \circ K = \text{ID}_U$. We suppose also $U \subseteq U'$. We define $V' := \{A \in \text{sent}_V \mid U' \vdash A^M\}$. Clearly, we have an interpretation $M' : V' \rightarrow U'$ based on the same translation as M . We have:

$$\begin{aligned}
 U' \vdash B &\Rightarrow U' \vdash B^{KM} \\
 &\Rightarrow V' \vdash B^K
 \end{aligned}$$

Hence there is an interpretation K' based on the same translation as K such that $K' : U' \rightarrow V'$. We find that K' is a section, since $B^{K'M'}$ is only dependent on the underlying translations, and hence strictly identical to B^{KM} . \square

¹¹ The notion of *sentential congruence* is explained in ‘‘Five categories’’ section of Appendix 1.

We need a sufficient store of incomparable extensions of given finitely axiomatised theories. We remind the reader that a theory U tolerates or weakly interprets a theory V if, for some translation τ , the theory $U + V^\tau$ is consistent. Note that we take the identity axioms for Σ_V including $\exists x x = x$ to be part of V . The following theorem can probably be much improved, but it is what we need for the current application.

Theorem 12 *Suppose A and B are finitely axiomatized theories that tolerate S_2^1 . Then, there are finitely axiomatized theories $A^* \supseteq A$ and $B^* \supseteq B$, that are incomparable w.r.t. \triangleleft , i.e. $A^* \not\triangleright B^*$ and $B^* \not\triangleright A^*$.*

Proof Suppose τ witnesses that A tolerates S_2^1 and ν witnesses that B tolerates S_2^1 . We take $A' := A + (S_2^1)^\tau$ and $B' := B + (S_2^1)^\nu$. So there is an N based on τ such that $N : S_2^1 \rightarrow A'$ and there is an M based on ν such that $M : S_2^1 \rightarrow B'$. By the Gödel Fixed Point Lemma, we find R such that:

$$EA \vdash R \Leftrightarrow \left((A' + R^N) \triangleright (B' + \neg R^M) \right) \leq \left((B' + \neg R^M) \triangleright (A' + R^N) \right).$$

We take $A^* := A' + R^N$ and $B^* := B' + \neg R^M$. Suppose $A^* \triangleright B^*$. It follows that R or R^\perp . In case we have R , we find, by Σ_1 -completeness, that $A' \triangleright \perp$. Quod non. If we have R^\perp , it follows that $(B' + \neg R^M) \triangleright (A' + R^N)$, by the fixed point equation. By Σ_1 -completeness, we have $B' \triangleright \perp$. Quod non. We may conclude that $A^* \not\triangleright B^*$.

The proof that $B^* \not\triangleright A^*$ is similar. □

The following theorem gives the basic simple insight concerning the unsplitability of connected theories. For the definition of *connected* see the ‘‘Sums’’ section of Appendix 1.

Theorem 13 *Suppose U and V are incomparable w.r.t. local interpretability. Then $U \boxplus V$ is not connected. It follows that no connected theory W can be mutually locally interpretable with $U \boxplus V$.*

Proof Suppose that U and V are incomparable w.r.t. local interpretability and that $U \boxplus V$ is connected. Since $(U \boxplus V) \triangleright_{loc} (U \boxplus V)$, it follows, by connectedness, that either U locally interprets $U \boxplus V$ or V locally interprets $U \boxplus V$. Suppose U locally interprets $U \boxplus V$. Then, $U \triangleright_{loc} (U \boxplus V) \triangleright_{loc} V$. So, $U \triangleright_{loc} V$. Quod non. The assumption that V locally interprets $U \boxplus V$ leads similarly to a contradiction. □

We are now ready to prove the main result of this section.

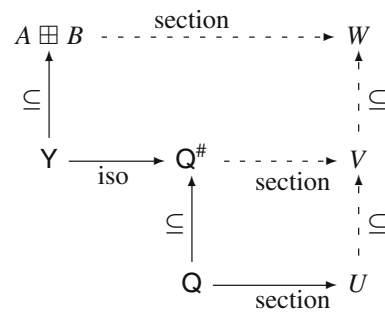
Theorem 14 *Q cannot be an INT₃-retract of a sequential theory.*

Proof We work in INT₃. Suppose U is sequential and Q is a retract of U . We derive a contradiction.

We have $Q \subseteq Q^\#$ and, hence by Theorem 11, we can find a theory $V \supseteq U$ such that $Q^\#$ is a retract of V . Since $Q^\#$ is synonymous with Y , the theory $Q^\#$ is a fortiori, sententially congruent to Y . It follows that Y is an INT₃-retract of V .

We can interpret S_2^1 in Q° , so Q° tolerates S_2^1 . We can extend CQC^2 to the weak set theory AS which interprets S_2^1 . So, CQC^2 tolerates S_2^1 . It follows that $CQC^2 \boxplus \mathbb{1}$ tolerates S_2^1 .¹² Let $A \supseteq Q^\circ$ and $B \supseteq CQC^2 \boxplus \mathbb{1}$ be the mutually incomparable theories promised by Theorem 12. By Theorem 11, we can find a theory $W \supseteq V$ such that $A \boxplus B$ is a retract of W .

It follows that $A \boxplus B$ is mutually locally interpretable with W . Moreover, A and B are incomparable w.r.t. local interpretability, since they are finitely axiomatized. This contradicts the result of Theorem 13.



We may conclude that Q is not a retract of a sequential theory. □

From our theorem, we immediately have that Q and PA^- are not sententially congruent.

We note that the only property we used in our proof of sequential theories is the fact that sequential theories are closed under theory-extension-in-the-same-language. Thus for any class \mathcal{X} of connected theories, such that \mathcal{X} is closed under theory-extension, we have that Q cannot be an INT₃-retract of \mathcal{X} .

8 Concluding remarks

The paper shows that the Pudlák Property and *being a poly-pair theory* are not preserved under mutual (faithful) interpretability. This provides two examples of good properties of theories that are not preserved under mutual (faithful) interpretability.¹³ It illustrates the usefulness of having more refined notions of sameness of theories.

Connectedness is preserved under mutual interpretability and even under mutual local interpretability. It is a notion of non-splitability. As we have seen the theory Q is splittable in

¹² I am using here that \boxplus is associative modulo synonymy.

¹³ A trivial example of a property not preserved under mutual interpretability is decidability. However, decidability is preserved under mutual faithful interpretability.

a sense, but this splittability is under the radar of the notion of connectedness, since \mathbf{Q} is connected. This discrepancy suggests that it might be interesting to explore more refined versions of connectedness that would exclude \mathbf{Q} but include \mathbf{PA}^- .

We note that our results in Sect. 7 imply that \mathbf{Q} is connected but has a non-connected extension, to wit (a theory synonymous to) a theory of the form $A \boxplus B$, where A and B are finitely axiomatized theories that are mutually incomparable w.r.t. relative interpretability. The natural class of sequential theories, however, is upwards closed under theory-extension. So, one may wonder if a more refined version of connectedness would have the property of being preserved under theory-extension.

Acknowledgements We thank Victor Pambuccian for his comments on the slides of a talk I gave on this subject. We are grateful to the anonymous referee for his perceptive comments and suggestions and for spotting some small errors. We thank Paul Taylor for the use of the Diagrams package.

Compliance with ethical standards

Conflict of interest The author declares that there are no conflicts of interest connected to publication of this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Basics

In this appendix, we provide detailed definitions of translations, interpretations and morphisms between interpretations.

Theories

Theories in this paper are one-sorted theories of first order predicate logic of finite relational signature. We take identity to be a logical constant. Our official signatures are relational, however, via the term-unwinding algorithm, we can also accommodate signatures with functions. For most purposes in the present paper, a theory can be identified with a deductively closed set of sentences of the given language. The exception is the few places where we use Rosser style arguments. We only do this in the context of finitely axiomatized theories. We assume that we employ the obvious axiomatizations in these cases.

Translations

Translations are the heart of our interpretations. In fact, they are often confused with interpretations, but we will not do

that officially. In practice it is often convenient to conflate an interpretation and its underlying translation.

To formulate the notion of translation, a number of subtleties and details concerning the choice and use of variables in the translations will be only sketched in a hand waving way.

We define more-dimensional, one-piece relative translations without parameters. Let Σ and Θ be one-sorted signatures. A translation $\tau : \Sigma \rightarrow \Theta$ is given by a triple $\langle m, \delta, F \rangle$. Here $\delta(v_0, \dots, v_{m-1})$ will be the domain formula. The mapping F associates to each relation symbol R of Σ with arity n a formula $A(\mathbf{v}_0, \dots, \mathbf{v}_{n-1})$ of signature Θ . Here the \mathbf{v}_i are sequences of variables of length m . The \mathbf{v}_i and \mathbf{v}_j , for $i \neq j$ are disjoint.

We demand that predicate logic proves

$$F(R)(\mathbf{x}_0, \dots, \mathbf{x}_{n-1}) \rightarrow (\delta(\mathbf{x}_0) \wedge \dots \delta(\mathbf{x}_{n-1})).$$

Of course, given any candidate $F(R)$ not satisfying the restriction, we can obviously modify the formula to satisfy the restriction.

We translate Σ -formulas to Θ -formulas as follows.¹⁴

- $(R(x_0, \dots, x_{n-1}))^\tau := F(R)(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})$.
- The single variable x_i of the source language needs to have no obvious connection with the sequence of variables \mathbf{x}_i of the target language that represents it. We need some conventions to properly handle the association $x_i \mapsto \mathbf{x}_i$.¹⁵
- $(\cdot)^\tau$ commutes with the propositional connectives;
- $(\forall x A)^\tau := \forall \mathbf{x}(\delta(\mathbf{x}) \rightarrow A^\tau)$;
- $(\exists x A)^\tau := \exists \mathbf{x}(\delta(\mathbf{x}) \wedge A^\tau)$.

Here are some convenient conventions and notations.

- We write δ_τ for ‘the δ of τ ’ and F_τ for ‘the F of τ ’.
- We write R_τ for $F_\tau(R)$.
- We write $\mathbf{x} \in \delta$ for: $\delta(\mathbf{x})$.

¹⁴ We assume possible variable clashes resulting by the substitution of the \mathbf{x} ’s for the \mathbf{v} ’s to be resolved by α -conversion.

¹⁵ There are several ways of handling such conventions. First we can work with a fixed global association between the x_i and the \mathbf{x}_i . Secondly, we can make such an association local and carry it around as an extra argument of the translation. Thirdly, we can throw away the mechanism of using variable-names and work in a language that works with explicit links between places. Fourthly, we can sidestep the problem by working in many-sorted languages and, for every k , adding sequences of length k (of various sorts). This construction can be viewed as a representation of more dimensional interpretations as arrows in a Kleisli category. Regrettably, each way of proceeding needs some work and produces some awkwardness somewhere. We demand that the \mathbf{x}_i are fully disjoint when the x_i are different. In this paper, we will assume that these details are taken care of by one strategy or by another.

There are some natural operations on translations. The identity translation $\text{id} := \text{id}_\Theta$ is one-dimensional and it is defined by:

- $\delta_{\text{id}}(x) := (x = x)$,
- $R_{\text{id}}(\mathbf{x}) := R\mathbf{x}$.

We can compose relative translations as follows. Suppose τ is an m -dimensional translation from Σ to Θ , and ν is a k -dimensional translation from Θ to Ξ . We define the $m \times k$ -dimensional interpretation $\tau\nu$ or $\nu \circ \tau$ as follows.

- We suppose that with the variable x we associate under τ the sequence x_0, \dots, x_{m-1} and under ν we send x_i to \mathbf{x}_i .

$$\delta_{\tau\nu}(\mathbf{x}_0, \dots, \mathbf{x}_{m-1}) := (\delta_\nu(\mathbf{x}_0) \wedge \dots \wedge \delta_\nu(\mathbf{x}_{m-1}) \wedge (\delta_\tau(x))^v),$$

- Let R be n -ary. Suppose that under τ we associate with x_i the sequence $x_{i,0}, \dots, x_{i,m-1}$ and that under ν we associate with $x_{i,j}$ the sequence $\mathbf{x}_{i,j}$. We take:

$$\begin{aligned} R_{\tau\nu}(\mathbf{x}_{0,0}, \dots, \mathbf{x}_{n-1,m-1}) \\ = \delta_\tau(\mathbf{x}_{0,0}) \wedge \dots \wedge \delta_\tau(\mathbf{x}_{n-1,m-1}) \wedge \\ (R_\tau(x_{0,0}, \dots, x_{n-1,m-1}))^v. \end{aligned}$$

We can make a disjunctive interpretation as follows. Suppose τ and ν are translations from Σ to Θ . We assume that τ is k -dimensional and ν is m -dimensional. Let A be a Θ -sentence. We introduce a $\max(k, m)$ -dimensional interpretation $\tau \langle A \rangle \nu$.

We first ‘lift’ one of the interpretations by padding to get the dimensions equal. Suppose, e.g. that $k < m$. Then we define:

- $\delta_{\tau'}(\mathbf{z}\mathbf{x}) := \delta_\tau(\mathbf{x})$,
- $P_{\tau'}(\mathbf{x}_0\mathbf{z}_0, \dots, \mathbf{x}_{n-1}\mathbf{z}_{n-1}) := P_\tau(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})$.

Here the dimension of the \mathbf{z} is $m - k$.

Suppose the results of the padding operation are τ' and ν' , where, of course, in case $k < m$, $\nu = \nu'$, etcetera. We define $\tau \langle A \rangle \nu$ as follows:

- $\delta_{\tau \langle A \rangle \nu}(\mathbf{x}) := ((A \wedge \delta_{\tau'}(\mathbf{x})) \vee (\neg A \wedge \delta_{\nu'}(\mathbf{x})))$.
- $R_{\tau \langle A \rangle \nu}(\mathbf{x}_0, \dots, \mathbf{x}_{n-1}) := ((A \wedge R_{\tau'}(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})) \vee (\neg A \wedge R_{\nu'}(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})))$.

Here the \mathbf{x} are $\max(k, m)$ -dimensional.

An m -dimensional translation τ preserves identity if

$$\mathbf{x} =_\tau \mathbf{y} := \bigwedge_{i < m} (\delta_\tau(x_i) \wedge \delta_\tau(y_i) \wedge x_i = y_i).$$

An m -dimensional translation τ is *unrelativized* if $\delta_\tau(\mathbf{x}) = \top$. An m -dimensional translation τ is *direct* if it is unrelativized and preserves identity. Note that all these properties are preserved by composition (modulo provable equivalence in predicate logic).

Consider a model \mathcal{M} with domain M of signature Σ and k -dimensional translation $\tau : \Sigma \rightarrow \Theta$. Suppose the τ -translations of the identity axioms, including $\exists x x = x$, are true in \mathcal{M} . Let $N := \{\mathbf{m} \in M^k \mid \mathcal{M} \models \delta_\tau \mathbf{m}\}$. Let E be the equivalence relation on N defined in \mathcal{M} by $=_\tau$. Then τ specifies an internal model \mathcal{N} of \mathcal{M} with domain N/E and with $\mathcal{N} \models R([\mathbf{m}_0]_E, \dots, [\mathbf{m}_{n-1}]_E)$ iff $\mathcal{M} \models R_\tau(\mathbf{m}_0, \dots, \mathbf{m}_{n-1})$. We will write $\tilde{\tau}(\mathcal{M})$ for the internal model of \mathcal{M} given by τ .

We treat the mapping $\tau, \mathcal{M} \mapsto \tilde{\tau}\mathcal{M}$ as a partial function that is defined precisely if the translations of the identity axioms are true in \mathcal{M} . Let Mod or $(\tilde{\cdot})$ be the function that maps τ to $\tilde{\tau}$. We have:

$$\text{Mod}(\tau \circ \rho)(\mathcal{M}) = (\text{Mod}(\rho) \circ \text{Mod}(\tau))(\mathcal{M}).$$

So, Mod behaves contravariantly.

Relative interpretations

A translation τ supports a *relative interpretation* of a theory U in a theory V , if, for all U -sentences A , we have $U \vdash A \Rightarrow V \vdash A^\tau$. Note that this automatically takes care of the theory of identity and assures us that δ_τ is inhabited. We will write $K = \langle U, \tau, V \rangle$ for the interpretation supported by τ . We write $K : U \rightarrow V$ for: K is an interpretation of the form $\langle U, \tau, V \rangle$. If M is an interpretation, τ_M will be its second component, so $M = \langle U, \tau_M, V \rangle$, for some U and V .

Par abus de langage, we write ‘ δ_K ’ for: δ_{τ_K} ; ‘ R_K ’ for: R_{τ_K} ; ‘ A^K ’ for: A^{τ_K} , etc. Here are the definitions of three central operations on interpretations.

- Suppose U has signature Σ . We define: $\text{ID}_U : U \rightarrow U$ is $\langle U, \text{id}_\Sigma, U \rangle$.
- Suppose $K : U \rightarrow V$ and $M : V \rightarrow W$. We define: $M \circ K : U \rightarrow W$ is $\langle U, \tau_M \circ \tau_K, W \rangle$.
- Suppose $K : U \rightarrow (V + A)$ and $M : U \rightarrow (V + \neg A)$. We define: $K \langle A \rangle M : U \rightarrow V$ is $\langle U, \tau_K \langle A \rangle \tau_M, V \rangle$.

It is easy to see that we indeed correctly defined interpretations between the theories specified.

Five categories

We do not automatically get a category of theories and interpretations from the machinery we built up until now. For example, $\text{ID}_U \circ \text{ID}_U$ will not be strictly speaking identical

with ID_U . We will obtain a category, when we divide out a suitable equivalence among interpretations. Below we will consider five kinds of equivalence that will give us five different categories. One important point of the categories is that isomorphism in each of them defines a salient notion of sameness of theories.

Provable equivalence of interpretations

Two interpretations are *provably equivalent* when the *target theory thinks they are the same*. Specifically, two interpretations $K, M : U \rightarrow V$ are provably equivalent if they have the same dimension, say m , and:

$$\begin{aligned} & - V \vdash \forall \mathbf{x} (\delta_K(\mathbf{x}) \leftrightarrow \delta_M(\mathbf{x})), \\ & - V \vdash \forall \mathbf{x}_0, \dots, \mathbf{x}_{n-1} \in \delta_K (R_K(\mathbf{x}_0, \dots, \mathbf{x}_{n-1}) \leftrightarrow R_M(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})). \end{aligned}$$

Modulo this identification, the operations identity and composition give rise to a category INT_0 , where the theories are objects and the interpretations arrows. Isomorphism in this category is *synonymy* or *definitional equivalence*. This is the strictest notion of identity between theories in the literature. It was first introduced by [de Bouvère \(1965a, b\)](#).

Let MOD be the category with as objects classes of models and as morphisms all functions between these classes. We define $Mod(U)$ as the class of all models of U . Suppose $K : U \rightarrow V$. Then, $Mod(K)$ is the function from $Mod(V)$ to $Mod(U)$ given by: $\mathcal{M} \mapsto \tilde{K}(\mathcal{M}) := \tau_{\tilde{K}}(\mathcal{M})$. It is clear that Mod is a *contravariant functor* from INT_0 to MOD .¹⁶

Maps between interpretations

For many applications provable equivalence is too strict. A better notions is provable isomorphism or *i-isomorphism*.

Consider $K, M : U \rightarrow V$. Suppose K is m -dimensional and M is k -dimensional. An *i-isomorphism* between interpretations $K, M : U \rightarrow V$ is given by a formula F with $m + k$ free variables in the language of V .¹⁷ We demand that V verifies that “ F is an isomorphism between K and M ”, or, equivalently, that, for each model \mathcal{M} of V , the function $F^{\mathcal{M}}$ is an isomorphism between $\tilde{K}(\mathcal{M})$ and $\tilde{M}(\mathcal{M})$.

We spell out the syntactical definition of an *i-isomorphism* $F : K \Rightarrow M$.

$$\begin{aligned} & - V \vdash \mathbf{x} F \mathbf{y} \rightarrow (\mathbf{x} \in \delta_K \wedge \mathbf{y} \in \delta_M). \\ & - V \vdash (\mathbf{x} =_K \mathbf{u} \wedge \mathbf{u} F \mathbf{v} \wedge \mathbf{v} =_M \mathbf{y}) \rightarrow \mathbf{x} F \mathbf{y}. \\ & - V \vdash \forall \mathbf{x} \in \delta_K \exists \mathbf{y} \in \delta_M \mathbf{x} F \mathbf{y}. \\ & - V \vdash \forall \mathbf{y} \in \delta_M \exists \mathbf{x} \in \delta_K \mathbf{x} F \mathbf{y}. \end{aligned}$$

¹⁶ Of course, there is a foundational issue with this definition. Let’s say that we work in Gödel–Bernays set theory to understand the definition.

¹⁷ The ‘*i*’ in ‘*i-isomorphism*’ stands for *interpretation*.

$$\begin{aligned} & - V \vdash (\mathbf{x}_0 F \mathbf{y}_0 \wedge \dots \wedge \mathbf{x}_{n-1} F \mathbf{y}_{n-1}) \rightarrow \\ & \quad (R_K(\mathbf{x}_0, \dots, \mathbf{x}_{n-1}) \leftrightarrow R_M(\mathbf{y}_0, \dots, \mathbf{y}_{n-1})). \end{aligned}$$

Here the last item includes identity in the role of R !

Two interpretations $K, M : U \rightarrow V$, are *i-isomorphic* iff there is an *i-isomorphism* between K and M . Wilfrid Hodges calls this notion: *homotopy*. See Hodges (1993, p. 222).

We can also define the notion of being *i-isomorphic* semantically. The interpretations $K, M : U \rightarrow V$, are *i-isomorphic* iff there is an F such that, for all V -models, \mathcal{M} , the relation $F^{\mathcal{M}}$ is an isomorphism between $\tilde{K}(\mathcal{M})$ and $\tilde{M}(\mathcal{M})$.

The default in this paper is that theories have finite signature: in this case we have a third characterization. The interpretations $K, M : U \rightarrow V$, are *i-isomorphic* iff, for every V -model \mathcal{M} , there is an \mathcal{M} -definable isomorphism between $\tilde{K}(\mathcal{M})$ and $\tilde{M}(\mathcal{M})$. This characterization follows by a simple compactness argument.

Clearly, if K and M are provably equivalent in the sense of the previous subsection, they will be *i-isomorphic*. The notion of *i-isomorphism* give rise to a category of interpretations modulo *i-isomorphism*. We call this category INT_1 .

Isomorphism in INT_1 is *bi-interpretability*. Bi-interpretability is a very good notion of sameness that preserves such diverse properties as finite axiomatizability and κ -categoricity.

Isomorphism

Our third notion of sameness of the basic list is that K and M are the same if, for all models \mathcal{M} of V , the internal models $\tilde{K}(\mathcal{M})$ and $\tilde{M}(\mathcal{M})$ are isomorphic. We will simply say that K and M are isomorphic. Clearly, *i-isomorphism* implies isomorphism. We call the associated category INT_2 . Isomorphism in INT_2 is *iso-congruence*.

Elementary equivalence

The fourth notion is to say that two interpretations K and M are the same if, for each \mathcal{M} , the internal models $\tilde{K}(\mathcal{M})$ and $\tilde{M}(\mathcal{M})$ are elementary equivalent. We will say that K and M are elementary equivalent.

By the completeness theorem, we easily see that this notion can be alternatively defined by saying that K is elementary equivalent to M iff, for all U -sentences A , we have $V \vdash A^K \leftrightarrow A^M$. It is easy to see that isomorphism implies elementary equivalence. We call the associated category INT_3 . Isomorphism in INT_3 is *elementary congruence* or *sentential congruence*.

Identity of source and target

Finally, we have the option of abstracting away from the specific identity of interpretations completely, simply counting any two interpretations $K, M : U \rightarrow V$ the same. The associated category is INT_4 . This is simply the structure of degrees of (global) interpretability DEG_{glob} . Isomorphism in INT_4 is *mutual interpretability*.

Global and local interpretability

We can view interpretability as a generalization of provability. When we take this stand point, we write:

- $U \triangleright V$ (or $V \triangleleft U$) for: U interprets V (or V is interpretable in U).
- $U \equiv V$ for: U and V are mutually interpretable.

A closely related notion is local interpretability. We define

- U locally interprets V or $U \triangleright_{\text{loc}} V$ iff, for every finitely axiomatize subtheory V_0 of V we have $U \triangleright V_0$.
- We write $V \triangleleft_{\text{loc}} U$ and $U \equiv_{\text{loc}} V$ with the obvious meanings.

If we want to stress the contrast between local and ordinary interpretability, we often call ordinary interpretability *global interpretability*. We will write $\triangleright_{\text{glob}}$, etcetera. The degrees of local interpretability are DEG_{loc} .

Example 1 Let \neq be the theory in the language of identity that says that there are precisely two elements. Let INF be the theory in the language of identity that has for every n an axiom saying ‘there are at least n elements’. Then $\neq \triangleright_{\text{loc}} \text{INF}$ but $\neq \not\triangleright_{\text{glob}} \text{INF}$

Weak interpretability or tolerance

We say that a theory U *weakly interprets* a theory V , or that U *tolerates* V if, for some interpretation $\tau : \Sigma_V \rightarrow \Sigma_U$, the theory $U + V^\tau$ is consistent. Here we take V to contain the axioms of identity including $\exists x x = x$. We note that:

- U tolerates V , iff for some consistent extension (in the same language) U' of U , we have $U' \triangleright V$.
- If U tolerates V and $V \triangleright W$, then U tolerates W .

Adding parameters

We can add parameters in the obvious way. An interpretation $K : U \rightarrow V$ with parameters will have a k -dimensional parameter domain α (given by a formula in k variables), where $V \vdash \exists \mathbf{x} \alpha \mathbf{x}$. We allow the extra variables \mathbf{x} to occur in the

translations of the U -formulas. We sometimes write $K^{\mathbf{x}}$ to make the dependence on the parameters visible. We write $A^{K,\mathbf{x}}$ for the K -translation of A for parameters \mathbf{x} .

The condition for K to be an interpretations changes into: $\vdash \forall \mathbf{x} (\alpha \mathbf{x} \rightarrow A^{K,\mathbf{x}})$, for all sentences A such that $U \vdash A$. Note that this automatically takes care of the axioms of identity and the non-emptiness of the domain.

Notions like *direct interpretation* are lifted in the obvious way to the case where we allow parameters.

We note that, in the presence of parameters, the functor \tilde{K} associates a class of models of U to a model of V .

Similar adaptations are needed to define i -isomorphisms with parameters.

Piecewise interpretability

The idea of piecewise interpretability is that we can build up the domain from a number of pieces that may or may not be of the same dimension and that may or may not overlap. The same object of the interpreting theory may occur in different roles posing as different objects of the interpreted theory. We will not develop piecewise interpretability here. We just given an example of how it works.

Suppose we have two pieces a and b . Let’s say that a is 1-dimensional and b is 2-dimensional. Suppose we have no parameters.

We want to translate $P(u_3, u_1)$. How are we going to do it? Well, we need to know in what pieces u_1 and u_3 are supposed to be. We need a function g as argument in the translation as an oracle that tells us precisely that. Suppose that, according to g , u_1 is in piece a and u_3 is in piece b . We note that g in combination with our formula $P(u_3, u_1)$ places b on the first argument place and a on the second argument place. So, F must tell us the translation for P when the first argument is b and the second argument is a . Let f be a function that associates pieces to argument places. Thus, we need $F(P, f)$, where $f(0) = b$ and $f(1) = a$. We may choose $F(P, f)$ be $A(u, v, w)$, where A is dependent on f . We may view $P(u_3, u_1)$ as given by P plus a function h that associates variables to argument places. Thus $(P(u_3, u_1))^{\tau, g} := F(P, g \circ h)(u_{30}, u_{31}, u_{10})$.

Here is the clause for the universal quantifier for the variable u_2 .

$$- (\forall u_2 A)^{\tau, g} := \left(\forall u_{20} \left(\delta_\tau^a(u_{20}) \rightarrow A^{\tau, g[u_2:a]} \right) \wedge \forall u_{20} \forall u_{21} \left(\delta_\tau^b(u_{20}, u_{21}) \rightarrow A^{\tau, g[u_2:b]} \right) \right),$$

where $g[u_2 : j]$ is the result of setting g at u_2 to j .

So we handle quantification over the different pieces simply by conjunction over the quantifications of each piece separately.

We can miraculously conjure any finite number of elements out of nothing, simply by creating sufficiently many pieces containing one element.

If the target theory V proves that we have at least two elements, we can always replace a piecewise interpretation by a many-dimensional piece-less one modulo i -isomorphism.

Sums

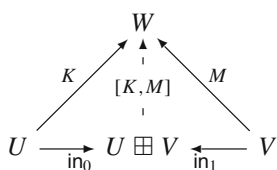
We define the operation \boxplus on theories as follows. The signature of $U \boxplus V$ is the disjoint union of the signatures of U and V , plus two new unary predicates Δ_0 and Δ_1 . The axioms of $U \boxplus V$ are:

- $P(x_0, \dots, x_{n-1}) \rightarrow \bigwedge_{i < n} \Delta_0(x_i)$, if P is derived from the signature of U ,
- $Q(y_0, \dots, y_{m-1}) \rightarrow \bigwedge_{j < m} \Delta_1(y_j)$, if Q is derived from the signature of V ,
- the axioms of U relativized to Δ_0 ,
- the axioms of V relativized to Δ_1 ,
- $\forall x(\Delta_0(x) \vee \Delta_1(x))$,
- $\forall x \neg(\Delta_0(x) \wedge \Delta_1(x))$.

We treat identity as outside of the signature here. We have the ordinary theory of identity.

We note that $U \boxplus V$ is synonymous with $V \boxplus U$ and $(U \boxplus V) \boxplus W$ is synonymous with $U \boxplus (V \boxplus W)$ and that both are synonymous with the ternary sum $\boxplus(U, V, W)$ which is defined in the obvious way using Δ_0, Δ_1 and Δ_2 .

We show that \boxplus is the sum in the categories INT_i for $1 \leq i \leq 4$ on the assumption that one of the theories U and V proves that there are at least two elements. We remind the reader of the sum diagram.



The arrow in_0 interprets U in $U \boxplus V$ by relativization to Δ_0 . We note that, by our conventions we should take $x =_{\text{in}_0} y$ iff $\Delta_0(x) \wedge \Delta_0(y) \wedge x = y$. The other predicate symbols do not need this addition. The definition of in_1 is similar.

Suppose $K : U \rightarrow W$ and $M : V \rightarrow W$. We suppose further that one of U and V proves that there are at least two elements. As a first step we make the domains of K and M in W disjoint and of the same dimension. We note that W proves that there are at least two elements. This follows from our assumption that at least one of U and V proves that there are at least two elements. Let the dimension of δ_K be k and the dimension of δ_M be

m . Suppose, e.g. $k \leq m$. We replace (x_0, \dots, x_{k-1}) in δ_K by $(z, z, w_0, \dots, w_{m-k-1}, x_0, \dots, x_{k-1})$ in $\delta_{K'}$, where z and the w_i can be arbitrary. We replace (y_0, \dots, y_{m-1}) in δ_K by $(z, z', y_0, \dots, y_{m-1})$, where z, z' can be arbitrary under the constraint that $z \neq z'$. Still under the assumption that $m \leq k$, we define $(z, z, w_0, \dots, w_{m-k-1}, x_0, \dots, x_{k-1}) =_{K'} (z', z', w'_0, \dots, w'_{m-k-1}, x'_0, \dots, x'_{k-1})$ iff $(x_0, \dots, x_{k-1}) =_K (x'_0, \dots, x'_{k-1})$. Etcetera. Let the newly obtained interpretations (considered as syntactical objects) be K' and M' . As is easily seen K' and M' have disjoint domains and the same dimensions and are equivalent to K , respectively M in each of INT_i , where $i = 1, 2, 3, 4$.

The interpretation $[K', M']$ is the obvious one where the new domain is the union of the domains of K' and M' . It is easy to see that, in each of the categories INT_i , for $i = 1, 2, 3, 4$, the arrow $[K', M']$ is unique with the desired property. Hence $U \boxplus V$ is indeed the sum in all of these theories.

As we have seen, the sum construction only works when at least one of the summands has provably two elements in its domain. If we have piecewise interpretations, we would not need this assumption. In this case the construction would become much simpler, since both padding and making disjoint become superfluous. There is an alternative construction of a sum $U \oplus V$ in Mycielski et al. (1990) or Stern (1989). This alternative construction is, for many purposes, more convenient. The reason that we use \boxplus in this paper is that in Theorem 3 we obtain synonymy for \boxplus . (The theories $U \oplus V$ and $U \boxplus V$ are bi-interpretable but not synonymous.)

We have the following basic theorem.

Theorem 15 Consider the theory $W := U \boxplus V$. Consider any formula $A\mathbf{x}\mathbf{y}$ in the language of W . Then, there are formulas $B_i\mathbf{x}$ in the language of U and formulas $C_j\mathbf{y}$ in the language of V , such that $A\mathbf{x}\mathbf{y}$ is equivalent to a boolean combination of $B_i^{\text{in}_0}\mathbf{x}$ and $C_j^{\text{in}_1}\mathbf{y}$ in the theory $W + \bigwedge_k \Delta_0(x_k) + \bigwedge_l \Delta_1(y_l)$.¹⁸

Proof The proof of the theorem is by a simple induction on A . □

An important notion that is defined in terms of the notion of *sum* is *connectedness*. We say that a theory W is *connected* if, for any theories U and V , if $(U \boxplus V) \triangleright_{\text{loc}} W$, then $U \triangleright_{\text{loc}} W$ or $V \triangleright_{\text{loc}} W$. The following fundamental theorem is due to Pudlák (1983). It was reproved with a markedly different proof by Stern (1989). For more context, see also Mycielski et al. (1990).

Theorem 16 Every sequential theory is connected.

The notion of *sequentiality* is introduced in Appendix 2.

¹⁸ Here we treat the variables x_k and y_l as parameters. They are not supposed to be universally generalized away.

Appendix 2: Pair theories and sequential theories

Pair theories and sequential theories are theories in which containers or data structures of a certain kind are present for all objects of a given theory. The presence of such containers provides many good properties for such theories. For example, recursively enumerable pair theories can be axiomatised by a scheme. See Visser (2012). Sequential theories are locally reflexive due to the presence of partial satisfaction predicates. We refer the reader to Visser (2013) for more information about poly-sequential theories.

Basic definitions

We consider the theory of non-surjective unordered pairing PAIR and adjunctive set theory AS. The language of both theories has just a binary predicate symbol \in (in addition to identity that is standardly available). The theory PAIR is axiomatised as follows:

- PAIR1. $\vdash \exists x \forall y y \notin x$
- PAIR2. $\vdash \exists z \forall u (u \in z \leftrightarrow (u = x \vee u = y))$

The theory AS is given by:

- AS1. $\vdash \exists x \forall y y \notin x$
- AS2. $\vdash \exists z \forall u (u \in z \leftrightarrow (u \in x \vee u = y))$

A theory is poly-sequential if it directly interprets AS. A theory is a poly-pair theory if it directly interprets PAIR. A theory is sequential if it directly interprets AS via a direct 1-dimensional interpretation. A theory is a pair theory if it directly interprets PAIR via a 1-dimensional interpretation.

Below we will mainly develop the poly-pair case. The proofs in the sequential case are entirely analogous.

We note that using the Kuratowski pairing we can obtain (non-extensional) ordered pairs in PAIR. Using iterated pairing we can obtain non-extensional sequences of length n , for each natural number n . Thus, we have a predicate SEQ_n and projection relations π_i for each $i < n$. We define

$$(x_0, \dots, x_{n-1}) \in_n y$$

$$:\leftrightarrow \exists z (\text{SEQ}_n(z) \wedge \pi_0(z, x_0) \wedge \dots \wedge \pi_{n-1}(z, x_{n-1}) \wedge z \in y).$$

One can show that:

- (i) $\vdash \exists x \forall y \quad y \notin_n x$
- (ii) $\vdash \exists z \forall \mathbf{u} \quad (\mathbf{u} \in_n z \leftrightarrow (\mathbf{u} = \mathbf{x} \vee \mathbf{u} = \mathbf{y}))$

Preservation under theory extension

Since direct interpretations are closed under composition, each theory that directly interprets a poly-pair theory is itself a poly-pair theory. Obviously, the identical embedding of a theory in an extension-in-the-same-language is direct. Ergo, being a poly-pair theory is preserved under extension-in-the-same-language. Similar remarks hold for poly-sequential theories.

Preservation under bi-interpretations

We show a slightly stronger result.

Theorem 17 *Let U be a poly-pair theory and suppose that V is a retraction in INT_1 of U . Then, V is a poly-pair theory.*

Proof To simplify the argument inessentially we ignore parameters.

Let $K : U \rightarrow V$ and $M : V \rightarrow U$ and let F be an i -isomorphism from ID_V to $K \circ M$ witnessing the retraction. We assume that K is k -dimensional and that M is m -dimensional. Let $(\cdot)^*$ be the interpretation of PAIR in U . Say $(\cdot)^*$ is ℓ -dimensional. We define a $k\ell$ -dimensional interpretation $(\cdot)^\dagger$ of PAIR in V as follows:

$$\mathbf{v} \in^\dagger \mathbf{u} :\leftrightarrow \exists \mathbf{w} \left(\mathbf{v} F \mathbf{w} \wedge \mathbf{w} \in_{km}^{\star K} \mathbf{u} \right).$$

Here:

- \mathbf{v} has length $k\ell$.
- \mathbf{w} has length $k^2\ell m$. It is of the form $\mathbf{w}_0, \dots, \mathbf{w}_{k\ell-1}$, where each \mathbf{w}_i has length km and $\mathbf{v}_i F \mathbf{w}_i$. The \mathbf{w}_i stand for elements of $\delta_{K \circ M}$. Note that the definition of an i -isomorphism forces the \mathbf{w}_i to be in $\delta_{K \circ M}$.
- \mathbf{u} has length $k\ell$. It stands for a sequence $\mathbf{u}_0, \dots, \mathbf{u}_{\ell-1}$, where the \mathbf{u}_j have length k . The \mathbf{u}_j stand for elements of δ_K . (In case they do not, we don't care. It is sufficient for our purposes that the 'correct' sequences \mathbf{u} provide all the unordered pairs we want.)
- The $\in_{km}^{\star K}$ lives inside K . Here its first component is $km\ell$ dimensional and its second component ℓ -dimensional. Looking at it from the outer level of ID_V , the first component acquires dimension $k^2\ell m$ and its second component acquires dimension $k\ell$.

A moment's reflection shows that our definition indeed gives us an interpretation of PAIR in V . We note that literally the same construction works for AS. □

Elimination of parameters

We show that we can eliminate the parameters from a witness of the property of being a pair theory.

Theorem 18 *Suppose U is a poly-pair theory. Then, U is a poly-pair theory via a direct interpretation without parameters.*

Proof Consider any poly-pair theory U . Let $(\cdot)^*$ be the witnessing interpretation of PAIR in U . Suppose $(\cdot)^*$ is k -dimensional and that M has an ℓ -dimensional parameter domain α . Let $ak + b = k + \ell$, where a and b are natural numbers and $b < k$. We define a parameter-free interpretation $(\cdot)^\dagger$ of dimension $k + \ell$ as follows.

$$\mathbf{x} \in^\dagger (\mathbf{p}, \mathbf{y}) :\Leftrightarrow \exists \mathbf{z}(\mathbf{x}, \mathbf{z}) \in_{a+1}^{\star \mathbf{p}} \mathbf{y}.$$

Here \mathbf{x} has length $k + \ell$, i.e. $ak + b$. The block \mathbf{z} is just padding and has length $k - b$. Thus, the length of (\mathbf{x}, \mathbf{z}) is $(a+1)k$. Finally, the length of \mathbf{y} is k and, therefore, the length of (\mathbf{p}, \mathbf{y}) is $k + \ell$.

We can easily verify that \in^\dagger yields an interpretation of PAIR. The argument in the case of AS is very similar. \square

Appendix 3: The converse of the Pudlák property fails

The following theorem is one of these utterly strange cases where the fact proven seems totally obvious, but where one still has to work to obtain the desired result. It would be interesting to prove the theorem using the quantifier elimination for EQ.

Theorem 19 *Let U and V be theories of finite signature. We suppose that U proves that there are at least two elements.¹⁹ Suppose V includes the theory of linear order for, say, $<$. Then any interpretation $K : V \rightarrow (U \boxplus \text{EQ})$ (with parameters) is i -isomorphic in $U \boxplus \text{EQ}$ to an interpretation with parameters $K^\star : V \rightarrow (U \boxplus \text{EQ})$, where both the parameter domain and the object domain consist of sequences from the U -domain Δ_0 .*

Proof Consider any model $\mathcal{M} \boxplus \mathcal{I}$ of $U \boxplus \text{EQ}$. We work in \mathcal{M} . We have the U -domain Δ_0 and the EQ-domain Δ_1 . We use a, b, c, \dots to range over Δ_0 , we use x, y, z, \dots to range over Δ_1 and we use u, v, w, \dots to range over the mixed domain.

Suppose that K is an m -dimensional interpretation and that K has a parameter domain that is ℓ -dimensional. Let s be the smallest number such that $2^s \geq \ell + m + 2$.

We define, for any n , a formula $G_n(\mathbf{a}, \mathbf{u})$, where the length of \mathbf{a} is sn and the length of \mathbf{u} is n . Here we intend to only consider G_n for $n \leq \ell + m$. The idea behind the formula G_n is that it represents the relation ‘ \mathbf{a} mimicks \mathbf{u} ’.

¹⁹ This assumption becomes superfluous when we allow piece-wise interpretations.

– $\mathbf{a} G_n \mathbf{u}$ iff

- (i) whenever u_i is in Δ_0 , then $a_{si} = a_{si+1} = \dots = a_{si+s-1} = u_i$;
- (ii) whenever u_i is in Δ_1 , then there are j and j' such that $j < j' < s$ and $a_{si+j} \neq a_{si+j'}$;
- (iii) whenever u_i and u_j are in Δ_1 , then:

$$u_i = u_j \quad \text{iff} \quad (a_{si}, \dots, a_{si+s-1}) = (a_{sj}, \dots, a_{sj+s-1}).$$

We define a translation τ^\star .

- $\alpha_{\tau^\star}(\mathbf{a}) :\Leftrightarrow \exists \mathbf{u}(\mathbf{a} G_\ell \mathbf{u} \wedge \alpha_{\tau_K}(\mathbf{u}))$,
- $\delta_{\tau^\star}^{\mathbf{a}}(\mathbf{b}) :\Leftrightarrow \exists \mathbf{u} \exists \mathbf{v}(\mathbf{a} \mathbf{b} G_{\ell+m} \mathbf{u} \mathbf{v} \wedge \delta_{\tau_K}^{\mathbf{u}}(\mathbf{v}))$,
- $P_{\tau^\star}^{\mathbf{a}}(\mathbf{b}_0, \dots, \mathbf{b}_{p-1}) :\Leftrightarrow \exists \mathbf{u} \exists \mathbf{v}_0 \dots \exists \mathbf{v}_{p-1} (\mathbf{a} \mathbf{b}_0 G_{\ell+m} \mathbf{u} \mathbf{v}_0 \wedge \dots \wedge \mathbf{a} \mathbf{b}_{p-1} G_{\ell+m} \mathbf{u} \mathbf{v}_{p-1} \wedge P_{\tau_K}(\mathbf{v}_0, \dots, \mathbf{v}_{p-1}))$.

We define the relation R between the parameters and the corresponding i -isomorphism F as follows:

- $\mathbf{a} R \mathbf{u} :\Leftrightarrow \mathbf{a} G_\ell \mathbf{u} \wedge \alpha_{\tau_K}(\mathbf{u})$.
- $\mathbf{b} F^{\mathbf{a}, \mathbf{u}} \mathbf{v} :\Leftrightarrow \mathbf{a} \mathbf{b} G_{\ell+m} \mathbf{u} \mathbf{v} \wedge \delta_{\tau_K}^{\mathbf{u}}(\mathbf{v})$.

Our first order of business is to show that R is total and surjective between α_{τ^\star} and α_{τ_K} . That it is total is immediate from the definition of α_{τ^\star} . We show that it is surjective. Consider any \mathbf{u} in α_{τ^\star} . To find a G_ℓ -counterpart \mathbf{a} , we have only to match the pattern of identity versus non-identity between the Δ_1 -elements of \mathbf{u} . To do that we need at most ℓ different sequences \mathbf{c} of length s that represent Δ_1 -elements—in the most demanding case the elements of \mathbf{u} are all in Δ_1 and are all different. Since Δ_1 contains at least two elements, we have at least 2^s sequences of length s . (We note that we do not need to introduce extra parameters for the two elements, since we are only interested in sameness and difference of these sequences.) Since the constant sequences are reserved for representing elements from Δ_0 , we need that $2^s \geq \ell + 2$. By our choice of s this is true.

Suppose we fix parameters \mathbf{a} and \mathbf{u} such that $\mathbf{a} R \mathbf{u}$. We want to show that $F^{\mathbf{a}, \mathbf{u}}$ is a bijection between $\delta_{\tau^\star}^{\mathbf{a}}$ and $\delta_{\tau_K}^{\mathbf{u}}$ modulo the respective identities of $\tau^{\star, \mathbf{a}}$ and $\tau_K^{\mathbf{u}}$.

We first prove that $F^{\mathbf{a}, \mathbf{u}}$ is total. Consider \mathbf{b} in $\delta_{\tau^\star}^{\mathbf{a}}$. By definition, there are \mathbf{u}' and \mathbf{v}' such that $\mathbf{a} \mathbf{b} G_{\ell+m} \mathbf{u}' \mathbf{v}'$. There is an automorphism σ of \mathcal{I} that maps Δ_1 -elements of \mathbf{u}' to the corresponding elements of \mathbf{u} . Take $\mathbf{v} := \sigma(\mathbf{v}')$. It follows that $\mathbf{b} F^{\mathbf{a}, \mathbf{u}} \mathbf{v}$.

We prove that $F^{\mathbf{a}, \mathbf{u}}$ is surjective. Consider $\mathbf{v} \in \delta_{\tau_K}^{\mathbf{u}}$. We need to find a \mathbf{b} such that $\mathbf{a} \mathbf{b} G_{\ell+m} \mathbf{u} \mathbf{v}$. For we need the sameness-difference pattern of the Δ_1 -elements of $\mathbf{a} \mathbf{b}$ with

the given pattern of the Δ_1 -elements of \mathbf{uv} . Since we have chosen $2^s \geq \ell + m + 2$ we can always do this.

We prove that $F^{a,\mathbf{u}}$ is functional. Suppose $\mathbf{b} =_{\tau^*}^a \mathbf{b}'$ and $\mathbf{b} F^{a,\mathbf{u}} \mathbf{v}$ and $\mathbf{b}' F^{a,\mathbf{u}} \mathbf{v}'$. We note that, for some \mathbf{w} and \mathbf{w}' , we have $\mathbf{ab} G_{\ell+m} \mathbf{uw}$ and $\mathbf{ab}' G_{\ell+m} \mathbf{uw}'$ and $\mathbf{w} =_{\tau_K}^u \mathbf{w}'$. We claim that $\mathbf{v} =_{\tau_K}^u \mathbf{w}$ and, similarly, that $\mathbf{v}' =_{\tau_K}^u \mathbf{w}'$. Assuming the claim, we have $\mathbf{v} =_{\tau_K}^u \mathbf{w} =_{\tau_K}^u \mathbf{w}' =_{\tau_K}^u \mathbf{v}'$ and we are done. We note that \mathbf{uv} and \mathbf{uw} have the same sameness-difference pattern on the Δ_1 elements.

We prove the claim. Suppose, to get a contradiction that $\mathbf{v} \neq_{\tau_K}^u \mathbf{w}$. It follows that either $\mathbf{v} <_{\tau_K}^u \mathbf{w}$, or $\mathbf{w} <_{\tau_K}^u \mathbf{v}$. We assume, e.g. that $\mathbf{v} <_{\tau_K}^u \mathbf{w}$. The other case is analogous.

Suppose Δ_1 (i.e. the domain of \mathcal{I}) is finite. There clearly is an automorphism σ of \mathcal{I} such that $\sigma(\mathbf{uv}) = \mathbf{uw}$, where we leave the Δ_0 -elements in place. We have:

$$\mathbf{v} <_{\tau_K}^u \sigma(\mathbf{v}) <_{\tau_K}^u \sigma^2(\mathbf{v}) <_{\tau_K}^u \dots$$

Since \mathcal{I} was supposed to be finite, this is impossible.

Suppose \mathcal{I} is infinite. We first assume that the Δ_1 -elements in \mathbf{v} and \mathbf{w} that are not in \mathbf{u} are disjoint. In this case there is a \mathcal{I} -automorphism σ such that $\sigma(\mathbf{uv}) = \mathbf{uw}$ and $\sigma(\mathbf{uw}) = \mathbf{uv}$. It follows that:

$$\mathbf{uw} = \sigma(\mathbf{uv}) <_{\tau_K}^u \sigma(\mathbf{uw}) = \mathbf{uv}.$$

A contradiction. So $\mathbf{v} =_{\tau_K}^u \mathbf{w}$. In case the Δ_1 -elements in \mathbf{v} and \mathbf{w} that do not occur in \mathbf{u} are not fully disjoint, we find a \mathbf{z} such that $\mathbf{ab} G_{\ell+m} \mathbf{uz}$ and the Δ_1 -elements in \mathbf{z} that are not in \mathbf{u} are disjoint from all Δ_1 -elements in \mathbf{u} , \mathbf{v} and \mathbf{w} . We now use our earlier argument, to see that $\mathbf{v} =_{\tau_K}^u \mathbf{z} =_{\tau_K}^u \mathbf{w}$.

We prove that $F^{a,\mathbf{u}}$ is injective. Suppose $\mathbf{b} F^{a,\mathbf{u}} \mathbf{v}$, $\mathbf{b}' F^{a,\mathbf{u}} \mathbf{v}'$ and $\mathbf{v} =_{\tau_K}^u \mathbf{v}'$. Then, by the definition of $=_{\tau^*}^a$, we find: $\mathbf{b} =_{\tau^*}^a \mathbf{b}'$.

Finally we want to show that, if $\mathbf{b}_0 F^{a,\mathbf{u}} \mathbf{v}_0$ and ...and $\mathbf{b}_{p-1} F^{a,\mathbf{u}} \mathbf{v}_{p-1}$, then:

$$P_{\tau^*}^a(\mathbf{b}_0, \dots, \mathbf{b}_{p-1}) \text{ iff } P_{\tau_K}^u(\mathbf{v}_0, \dots, \mathbf{v}_{p-1}).$$

But this is immediate by the definition of $P_{\tau^*}^a$ and the fact that F is a bijection.

We define $K^* := \langle V, \tau^*, U \boxplus \text{EQ} \rangle$. It is easy to see that, by the Completeness Theorem, K^* satisfies the desiderata. \square

We are now ready to prove the Pudlák's Property for $U \boxplus \text{EQ}$, whenever U is sequential.

Theorem 20 *Suppose U is sequential. Then, we have both the Pudlák property and the strong Pudlák property for $U \boxplus \text{EQ}$.*

Proof We just consider the strong property. Suppose $N : S_2^1 \rightarrow U \boxplus \text{EQ}$. We can find $N^* : S_2^1 \rightarrow U \boxplus \text{EQ}$ that is

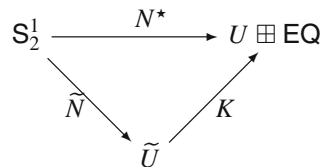
i -isomorphic to N such that the parameter domain α and the object domain δ of N^* are entirely in Δ_0 . Note that the definitions of the corresponding translations are not necessarily entirely in the U -part of the language.

Suppose the dimension of the parameter domain of N^* is k and the dimension of N^* is m . We expand the signature Σ_U of U with new symbols $\tilde{\alpha}, \tilde{\delta}, \tilde{Z}, \tilde{A}, \tilde{M}$. Here $\tilde{\alpha}$ is k -dimensional, \tilde{Z} is $(k + m)$ -dimensional, \tilde{S} is $(k + 2m)$ -dimensional, \tilde{A} is $(k + 3m)$ -dimensional, \tilde{M} is $(k + 3m)$ -dimensional. Let the resulting signature be $\tilde{\Sigma}_U$. We define a translation $\tilde{\tau}$ from the signature of arithmetic into $\tilde{\Sigma}_U$, by taking $\alpha_{\tilde{\tau}}(\mathbf{u}) := \tilde{\alpha}(\mathbf{u})$, $\delta_{\tilde{\tau}}(\mathbf{u}, \mathbf{v}) := \tilde{\delta}(\mathbf{u}, \mathbf{v})$, $Z_{\tilde{\tau}}(\mathbf{v}, \mathbf{u}) := \tilde{Z}(\mathbf{v}, \mathbf{u})$, etcetera.

Let $\tilde{U} := U + \forall \mathbf{u} (\tilde{\alpha}(\mathbf{u}) \rightarrow (S_2^1)^{\tilde{\tau}, \mathbf{u}})$. Here the axioms of S_2^1 are supposed to include the identity axioms. Let $\tilde{N} : S_2^1 \rightarrow \tilde{U}$ be the interpretation based on $\tilde{\tau}$.

We define a parameter-free interpretation $K : \tilde{U} \rightarrow (U \boxplus \text{EQ})$ by letting τ_K restricted to Σ_U be the translation corresponding to in_0 . We take $\tilde{\alpha}_{\tau_K}(\mathbf{v}) := \alpha_{N^*}(\mathbf{v})$, $\tilde{\delta}_{\tau_K}(\mathbf{u}, \mathbf{v}) := \delta_{N^*}(\mathbf{u}, \mathbf{v})$, $S_{\tau_K}(\mathbf{uv}) := S_{N^*}(\mathbf{u}, \mathbf{v})$, etcetera. It is easy to see that τ_K indeed supports the promised interpretation K .

It is easy to see that the following diagram commutes in INT_0 .



Clearly, \tilde{U} is a sequential theory. So, we have the strong Pudlák property for \tilde{N} . Thus, we have a parameter-free $N_0 : S_2^1 \rightarrow \tilde{U}$ with $N_0 \leq \tilde{N}$.

We find that $K \circ N_0$ is parameter-free and $(K \circ N_0) \leq (K \circ \tilde{N})$. On the other hand, $K \circ \tilde{N}$ is equal in the sense of INT_0 to N^* . Ergo, $(K \circ N_0) \leq N^*$. Since N^* is i -isomorphic to N , we find $(K \circ N_0) \leq N$. \square

References

Boolos GS, Burgess JP, Jeffrey RC (2002) Computability and logic, 4th edn. Cambridge University Press, Cambridge

Bezboruah A, Shepherdson JC (1976) Gödel's second incompleteness theorem for Q. J Symb Log 41(2):503–512

de Bouvère KL (1965a) Logical synonymy. Indag Math 27:622–629

de Bouvère KL (1965b) Synonymous theories. In: Addison JW, Henkin L, Tarski A (eds) Proceedings of the 1963 international symposium at the theory of models at Berkeley. North Holland, Amsterdam, pp 402–406

Friedman H (2007) Interpretations according to Tarski. This is one of the 2007 Tarski lectures at Berkeley. The lecture is available at <http://www.math.osu.edu/~friedman.8/pdf/Tarski1.052407>

Hodges W (1993) Model theory, vol 42. Encyclopedia of mathematics and its applications. Cambridge University Press, Cambridge

- Hájek P, Pudlák P (1993) Metamathematics of first-order arithmetic. Perspectives in mathematical logic. Springer, Berlin
- Jeřábek E (2012) Sequence encoding without induction. *Math Log Q* 58(3):244–248
- Jones JP, Shepherdson JC (1983) Variants of Robinson’s essentially undecidable theory R. *Archiv Math Logik Grundl* 23:61–64
- Kaye R (1991) Models of peano arithmetic. Oxford University Press, Oxford
- Mycielski J, Pudlák P, Stern AS (1990) A lattice of chapters of mathematics (interpretations between theorems), vol 426. *Memoirs of the American Mathematical Society (AMS)*, Providence
- Mostowski A, Tarski A (1949) Undecidability in the arithmetic of integers and in the theory of rings. *J Symb Log* 14:76
- Nelson E (1986) *Predicative arithmetic*. Princeton University Press, Princeton
- Pambuccian V (2008) The sum of irreducible fractions with consecutive denominators is never an integer in PA^- . *Notre Dame J Form Log* 49(4):425–429
- Pambuccian V (2014) The Green–Tao theorem on primes in arithmetical progressions in the positive cone of $\mathbb{Z}[X]$. *Elem Math* 69(1):30–32
- Pambuccian V (2015) Schatunowsky’s theorem, Bonse’s inequality, and Chebyshev’s theorem in weak fragments of Peano arithmetic. *Math Log Q* 61(3):230–235
- Pour-El MB, Kripke S (1967) Deduction-preserving “recursive isomorphisms” between theories. *Fundam Math* 61:141–163
- Pudlák P (1983) Some prime elements in the lattice of interpretability types. *Trans Am Math Soc* 280:255–275
- Pudlák P (1985) Cuts, consistency statements and interpretations. *J Symb Log* 50(2):423–441
- Robinson RM (1950) An essentially undecidable axiom system. *Proc Int Congr Math* 1:729–730
- Stern AS (1989) Sequential theories and infinite distributivity in the lattice of chapters. *J Symb Log* 54:190–206
- Švejdar V (2007) An interpretation of Robinson’s arithmetic in its Grzegorzczk’s weaker variant. *Fundam Inform* 81(1–3):347–354
- Tarski A, Mostowski A, Robinson RM (1953) *Undecidable theories*. North-Holland, Amsterdam
- Vaught RL (1962) On a theorem of Cobham concerning undecidable theories. In: Nagel E, Suppes P, Tarski A (eds) *Proceedings of the 1960 international congress on logic, methodology and philosophy of science*. Stanford University Press, Stanford, pp 14–25
- Visser A (2006) Prolegomena to the categorical study of interpretations. *Logic Group preprint series No. 249*. Utrecht University, Utrecht. <http://www.phil.uu.nl/preprints/lgps/>
- Visser A (2008) Pairs, sets and sequences in first order theories. *Arch Math Log* 47(4):299–326
- Visser A (2012) Vaught’s theorem on axiomatizability by a scheme. *Bull Symb Log* 18(3):382–402
- Visser A (2013) What is sequentiality? In: Cégielski P, Cornaros C, Dimitracopoulos C (eds) *New studies in weak arithmetics*, vol 211. CSLI lecture notes. CSLI Publications and Presses Universitaires du Pôle de Recherche et d’Enseignement Supérieur Paris-est, Stanford, pp 229–269
- Visser A (2014) Interpretability degrees of finitely axiomatized sequential theories. *Arch Math Log* 53(1–2):23–42
- Visser A (2014) Why the theory \mathbb{R} is special. In: Tennant N (ed) *Foundational adventures. Essays in honour of Harvey Friedman*. College Publishing Ltd, London, pp 7–23 (originally published online by Templeton Press in 2012). <http://foundationaladventures.com/>